

Міністерство освіти і науки України
Національний технічний університет України
“Київський політехнічний інститут імені Ігоря Сікорського”
Міністерство освіти і науки України
Національний технічний університет України
“Київський політехнічний інститут імені Ігоря Сікорського”

Кваліфікаційна наукова
праця на правах рукопису

Осауленко Вячеслав Миколайович

УДК 004.8

ДИСЕРТАЦІЯ

**Моделі біологічних нейронних мереж для просторово-часової
асоціативної пам’яті**

05.13.23 – системи та засоби штучного інтелекту

Подається на здобуття наукового ступеня кандидата технічних наук

Дисертація містить результати власних досліджень. Використання ідей,
результатів і текстів інших авторів мають посилання на відповідне джерело

В.М.Осауленко

Науковий керівник д.ф.-м.н., професор О.С. Макаренко

Київ – 2019

АНОТАЦІЯ

Осауленко В.М. Моделі біологічних нейронних мереж для просторово-часової асоціативної пам'яті. - Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.23 – системи та засоби штучного інтелекту – Національний технічний університет України “Київський політехнічний інститут імені Ігоря Сікорського” МОН України, Київ, 2019 р.

У цій роботі представлені біологічно подібні моделі просторово-часової асоціативної пам'яті, засновані на останніх ідеях та відкриттях з нейронауки. Архітектури більшості штучних нейронних мереж, наприклад, згорткових нейронних мереж, базуються на біологічних ідеях, що датуються 60-80-х роками минулого століття. Однак, відтоді наше розуміння нейронних обчислень суттєво змінилося, і нові результати повинні бути формалізовані, досліджені та включені в нові алгоритми.

Термін у темі дисертації "просторово-часова асоціативна пам'ять" не є поширеним і бере своє походження від терміну "просторово-часова нейронна динаміка". Просторова пам'ять означає збережений патерн нейронної активації в будь-який момент часу. Просторово-часова пам'ять враховує зміну патерну активації в часі. Моделі біологічних нейронних мереж означають біологічно натхненні та біологічно подібні моделі.

Робота починається з широкого, але стислого огляду біологічних та штучних нейронних мереж (ШНМ). Коротко описані основні підходи трьох поколінь ШНМ: прості моделі МакКаллох-Піттса, сучасний підхід DeepLearning та спайкові нейронні мережі з окресленням їх основних проблем. Особливий інтерес приділяється обробці інформації в мозку. Коротко обговорюються теми нейронного кодування, мікроконнектому кори головного мозку та основних шляхів передачі інформації між областями. Зазначається, що однією з головних проблем є розуміння того, як відбувається навчання в одному нейроні, що

узгоджується з динамікою всієї мережі.

Перша модель представляє двовимірну модель двонаправленої пластичності в простій спайковій нейронній мережі. Модель включає в себе триpletне спайк-часово-залежне правило пластичності, правило гомеостатичної пластичності та сигнал підкріплення, що значно прискорює навчання та стабілізує зв'язки. Модель ґрунтується на тому, що молекулярні механізми в дендритних шипиках працюють на різних часових масштабах. Синапс характеризується двома змінними, швидкою та повільною вагою. Швидка, як і звичайна, визначає збудження від входу і змінюється через хебоподібні правила навчання. Повільна вага впливає на швидку, як і сама зазнає впливу від швидкої тільки на різних масштабах часу. Повільна вага змінюється від гомеостатичної пластичності і може швидко змінюватися від сигналу підкріплення. Показано, що модель забезпечує вивчення стійких з часом асоціативних зв'язків, і що важливо, новий досвід зберігає раніше навчені ваги. Модель має можливу біологічну інтерпретацію, де швидка вага представляє кількість та ефективність AMPA рецепторів, а повільна вага являє собою повільні молекулярні механізми, можливо, на основі CAMKII, та поверхневі мембранні ефекти, що визначають стабілізацію рецепторів AMPA. Однак, стверджується, що спайкові нейромережі все ще не можуть бути застосовані на практиці, зокрема, в задачах робототехніки, тому решта роботи присвячена більш математично абстрактним моделям, які легше аналізувати.

Наступна модель враховує дендритні обчислення, зокрема, синаптичну кластеризацію. Нещодавно було показано, що синапси з активних вхідних нейронів групуються локально на дендритному сегменті, так що пізніша реактивація може викликати супралінійне сумування та дендритний спайк. Тобто, відбувається детектування збігів, що відображається в моделі сігма-пай нейрона, яка в цій роботі використовується для побудови асоціативної пам'яті. Раніше вивчали ємність пам'яті одного нейрона з дендритами, але ця робота йде далі і досліджує ємність пам'яті мережі з нейронами сігма-пай. Теоретичний та об-

числювальний аналіз показує, що мультиплікативна активація для розрідженої активації популяції нейронів дає значно більшу ємність пам'яті. Вона знижує помилкову активацію нейрона, відому як cross-talk, завдяки відстеженню кореляцій вищого порядку, тобто виявленню збігів у вхідній активності. Порівняння зі стандартною моделлю Вілшоу показує, що представлена модель має більшу синаптичну ємність, але меншу інформаційну ємність: міри ємності пам'яті, що відображають кількість синапсів та кількість біт на синапс відповідно. Для апаратної реалізації синаптична ємність важливіша, оскільки вона відображає фактичну кількість з'єднань, необхідних для зберігання інформації. Цей результат узгоджується з біологічних фактів, що пов'язані з пам'яттю нейронні мережі в мозку розріджено активні, а нейрони мають розгалужене дендритне дерево. Цікаво, що з еволюційної точки зору більш розумні тварини мають більші дендритні дерева. У світлі представленої роботи стає зрозуміло, що більш складний нейрон збільшує ємність пам'яті мережі в цілому.

Результат більшої ємності з розріджено активною мережею та нейроном сігма-пай був використаний для завдання прогнозування послідовностей. Подібно до мереж із затримкою часу, контекст був побудований як активність з останніх T часових кроків. Представлена модель порівнювалася з моделлю ієрархічної темпоральної пам'яті (НТМ), яка також використовує розріджене представлення та модель нейронів з дендритами. Аналіз показує порівняно більшу на один порядок ємність пам'яті переходів в послідовності при однаковій кількості зв'язків. Така різниця виникає через мультиплікативну модель сігма-пай нейрона та спрощену архітектуру без мініколонок. Однак у моделі є недолік – використання контексту як останніх кроків часу T запобігає передбаченню більш віддалених в часі елементів послідовності.

Остання модель стосується часової інтеграції або представлення послідовності у вигляді розрідженого бінарного вектора. Вже встановлено, що пам'ять зберігається (алокується від англ. *allocation*) в нейронах з підвищеною здатністю до активації (*excitability*) на масштабі часу годин. Ця робота пропонує гіпотезу,

що здатність до активації важлива також на менших масштабах часу порядку секунд. Щоб перевірити це, стандартний алгоритм випадкової проекції з нелінійністю *k-winners-take-all* було змінено з урахуванням здатності нейрона до активації. Обчислювальні експерименти показують, що додавання здатності до активації додає короткочасну пам'ять і відтворює деякі ефекти представлення послідовності в біологічних нейронних мережах. Ці ефекти включають: схожі послідовності мають схожі представлення, елементи в одній послідовності представляються схожим чином, початок послідовності представлений схоже до представлення всієї послідовності. Використано схожість косинусів для порівняння різних представлень. Також, модель може формувати представлення, яке може використовуватися як контекст для попередньої моделі прогнозування послідовностей. Згідно з результатами, механізм алокації пам'яті в нейрони зі збільшеною здатністю до активації важливий також на масштабі часу секунд, що повинно далі бути підтверджено в біологічних експериментах.

Представлені моделі були досліджені на можливість практичного застосування для розпізнавання образів у робототехніці. Така задача вимагає навчання на малій кількості прикладів, в режимі реального часу, та змінною кількістю класів; все це обмежує використання традиційних згорткових нейронних мереж. Природний спосіб спробувати біологічно подібні алгоритми для типової біологічної задачі показав низькі результати точності, які залежать від багатьох рукотворних (*hand-crafted*) розв'язків. Тим не менш, в результаті стажування в лабораторії робототехніки Logia деякі результати цього рукопису були використані для їх роботи. Крім того, аналіз ємності пам'яті асоціації з нейроном сігма-пай був включений до курсу Київського політехнічного інституту Ігоря Сікорського з Нейронних мереж.

Загалом, новизна роботи може бути зведена до:

- досліджено ємність асоціативної пам'яті з нейроном сігма-пай, що відображає детектування збігів дендритами;
- показано, що врахування здатності нейрона до активації призводить до

біологічних ефектів представлення послідовностей і має корисні обчислювальні властивості;

- досліджено правило двовимірної двонаправленої пластичності в спайковій нейромережі;

- покращена ємність пам'яті переходів в послідовності в мережі з сігма-пай нейроном і розріджено розподіленим представленням даних.

Сподіваюся, розв'язані задачі в цій роботі дадуть хоча б невеликий внесок у головну мету – створення загальної теорії обробки інформації в мозку.

Ключові слова: асоціативна пам'ять, дендритні обчислення, передбачення послідовностей, розпізнавання послідовностей, сігма-пай нейрон, розріджено розподілене представлення

SUMMARY

Osaulenko V.M. Neural networks for spatio-temporal association memory. - The manuscript.

Thesis for a candidate degree in specialty 05.13.23 - systems and applications of artificial intelligence. - National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute Kyiv, 2019.

This work presents biologically plausible models of spatio-temporal association memory based on recent ideas and discoveries from neuroscience. The architectures of most artificial neural networks, like convolutional neural networks, are based on biological ideas dated back to 60-80s of the previous century. However, since then our understanding of neural computation changed dramatically and new evidence should be formalized, investigated and incorporated into new algorithms.

The term in the topic "spatio-temporal association memory" is not common and takes its origin from the term "spatio-temporal neural dynamics". Spatial memory means the stored pattern of neural activation at any given moment of time. Spatio-temporal memory relates to the pattern of activity spread in time. Models of biological neural networks mean biologically inspired and biologically plausible models.

The manuscript starts with a broad but concise overview of the biological and artificial neural networks (ANN). It briefly describes main approaches from three generations of ANN: simple McCulloch-Pitts models, modern DeepLearning approach, and spiking neural networks with outlining their main problems. Particular interest is devoted to information processing in the brain. The topics of neural coding, neocortex microcircuitry and large scale brain connectivity with main pathways of information transmission briefly discussed. It states that one of the main problems is to understand how learning in a single neuron occurs so that meaningful dynamics of the whole network arises.

The first model presents a two-dimensional bidirectional plasticity model in a simple spiking neural network. The model includes triplet spike-time-dependent plasticity rule, homeostatic plasticity rule and reinforcement signal that significantly speeds up learning and stabilizes the weights. It is based on the findings that molecular mechanisms in the dendritic spines work on different time scales and is similar to eligibility traces. The synapse is characterized by the two variables, the fast and the slow weight. The fast is like usual weight and determines the excitation from incoming input and is changing from the Hebbian rules. The slow weight influences the fast and gets influenced by fast in a bidirectional manner with different time scales. The slow weight is changing from the homeostatic plasticity and can rapidly be changed from reinforcement signal. The model shows learning of associative connections that are stable with time, and importantly, new experience does not significantly disrupt previously learned weights. The model has a possible biological interpretation, where the fast weight represents the number and efficiency of AMPA channels and slow weight represents slow molecular mechanisms, possibly CAMKII based, and surface effects on a membrane that determine stabilization of AMPA receptors. However, it is argued that spiking neural networks are still inapplicable for a practical application like for robotics, so the rest of the work continues with more mathematically abstract, but computationally trackable models.

The next model takes into account dendritic computation, particularly, synaptic

clustering. Recently, it was shown that synapses from active input neurons group locally on the dendritic branch, so that later reactivation can cause supralinear summation and dendritic spike. Such coincident detection nature is reflected in a sigma-pi neuron model that in this work used to construct association memory. Previously, the memory capacity of a single neuron with dendrites was studied, but this work goes further and investigates the network memory capacity with such neurons. The theoretical and computational analysis shows that multiplicative activation for sparsely active neuron population gives much higher memory capacity. It reduces the neuron false activation, known as cross-talk, due to higher-order correlation tracking or coincidence detection in the incoming activity. Comparison to standard Willshaw model shows that the presented model has higher synaptic capacity but lower information capacity, measures of memory capacity that reflect the number of synapses and number of bits per synapse respectively. For hardware implementation, the synaptic capacity is more important since it reflects the actual number of connections needed for storing information. This result follows biological facts that memory-related neural networks in the brain are sparsely active and neurons have an extensive dendritic tree. Interestingly, from the evolutionary perspective, more intelligent animals have more elaborate dendritic trees. In light of the presented work, it becomes clear that more complex neuron gives higher memory capacity of the network overall.

The result of higher capacity with sparsely active network and sigma-pi neuron was used for sequence prediction task. Similarly to time-delay networks, the context was constructed as activity from last T time steps. The presented model was compared to hierarchical temporal memory (HTM) model that also uses a sparse distributed representation and neuron model with dendrites. The analysis shows higher transition memory capacity by one order of magnitude with the same number of connections. Such difference arises because of the multiplicative neuron sigma-pi model and simplified architecture without minicolumns. However, the model has a disadvantage that using the context as the last T time steps prevents predicting more distant

relationships.

The final model deals with the temporal integration or sequence representation as sparse distributed binary vector. It is based on the established result that memory is allocated to neurons with increased excitability within hours time scale. This work proposes the hypothesis that excitability is important on smaller (seconds) time scale. To check this, the standard algorithm of random projections with k-winners-take-all nonlinearity was modified to include neuronal activity-dependent excitability. The computational simulation shows that addition of excitability creates short term memory and recreates some of the sequence representation effects in biological neural networks. These effects include: similar sequences represented similarly, events in the same sequence represented similarly, beginning of a sequence represented similarly as the whole sequence. Cosine similarity was used to compare different representations. The model can form representation that could be used as a context for the sequence prediction model. Results suggest that the mechanism of memory allocation to excitable neurons is important even for seconds time scale and should be confirmed on biological experiments.

Presented models were investigated for possible practical application for object recognitions in robotics. The task requires learning on a few examples, online, with a non-fixed number of classes; all these restrict from using traditional convolutional neural networks. The natural way to try biologically inspired algorithms for biologically relevant recognition showed poor accuracy results that depend on many handcrafted solutions. As a result of the internship to robotics laboratory Loria, some results of this manuscript were used for their work. Also, analysis of association memory capacity with sigma-pi neuron was incorporated to Igor Sikorsky Kyiv polytechnic institute course on Neural networks.

Overall, the novelty of the work can be summarized into:

- investigated the memory capacity of the association memory with a sigma-pi neuron that reflects coincident detection by dendrites
- showed that model with neural excitability leads to biological effects of sequence

representation and have useful computational properties

- investigated two-dimensional bidirectional plasticity rule in the spiking neural network

- improved sequence transition capacity in the network with sigma-pi neuron and sparse distributed representation.

Hopefully, this work provides at least a small contribution to the main long-awaited goal of creating the general theory of information processing in the brain.

Keywords: associative memory, dendritic computation, sequence prediction, sequence recognition, sigma-pi neuron, sparse distributed representation

Список опублікованих праць за темою дисертації

1. Osaulenko, V., Girau, B., Makarenko, O. Henaff, P. Increasing Capacity of Association Memory by Means of Synaptic Clustering. Neural Process. Lett. (2019), pp 1–14 doi:10.1007/s11063-019-10051-7 Особистий внесок: розробка моделі, отримання теоретичних та обчислювальних результатів. (Іноземне видання, реферується в SCOPUS)
2. Osaulenko V.M. On the quest to recreate intelligence . Review of Exploring robotic minds : Actions , symbols , and consciousness as self-organizing dynamic. Cognitive Systems Research, 2018, 51, 56–60. <https://doi.org/10.1016/j.cogsys.2018.04.008> (Іноземне видання, реферується в WebOfScience)
3. Osaulenko V.M Simple model for sequence prediction based on dendritic spatiotemporal integration System Research & Information Technologies, 2018, № 4, 133-141, DOI:10.20535/SRIT.2308-8893.2018.4.11 (Фахове видання, індексується наукометричними базами DOAJ, Index Copernicus, Google Scholar, PBN)
4. В. М. Осауленко Модель інтелектуального агента в одномуірному дискретному світі ISSN 1999-9941, “Інформаційні технології та комп’ютерна інженерія”, 2018, № 3, 30-36, DOI: <https://doi.org/10.31649/1999-9941-2018->

43-3-30-36. (Фахове видання, індексується Google Scholar)

5. Осауленко В.М. Двовимірна модель навчання у спайкових нейронних мережах з гомеостазом та навчанням з підкріпленням, System Research & Information Technologies, 2017, № 2, 130-140, DOI: 10.20535/SRIT.2308-8893. 2017.1.12 (Фахове видання, індексується наукометричними базами DOAJ, Index Copernicus, Google Scholar, PBN)
6. Osaulenko V.M. Testing simple neuron models with dendrites for sparse binary image representation, Штучний інтелект, 2017, № 2, С. 101-108 (Фахове видання, індексується Google Scholar)
7. Makarenko O.S., Osaulenko V.M. Application of cellular automates in some models of artificial intelligence. Конференція IEEE SAIC-2018, 2018, 1-4. 10.1109/SAIC.2018.8516837
8. Osaulenko V.M. Learning of higher-order correlations improves capacity of an association memory. Nonlinear analysis and applications 4rd International Conference of memory of corresponding member of National Academy of Sciences of Ukraine Valeriy Sergeevich MELNIK (1952-2007) NTUU KPI, Kyiv, Ukraine, April 04-06, 2018, p.76.
9. Osaulenko V.M. Model for sequence prediction based on dendritic spatio-temporal integration. 20-th International conference on System Analysis and Information Technology SAIT 2018, May 21-23, 2018, pp.36-37. Institute for Applied System Analysis of National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute Kyiv, Ukraine
10. Osaulenko V.M. Dendritic nonlinearity provides high capacity sequence memory FENS FORUM of neuroscience 2018, p.368, Berlin
11. Osaulenko V.M. New ideas to test for artificial neural networks inspired from neuroscience Міжнародна наукова конференція «Сучасна інформатика:

проблеми, досягнення та перспективи розвитку» Інституту кібернетики імені В.М.Глушкова, 13-15 грудня 2017, с.82-83

12. Осауленко В.М. Тестування простих моделей з дендритами для розрізженого бінарного представлення зображень. XVII Міжнародна науково-технічна конференція Штучний інтелект та інтелектуальні системи. Artificial intelligence and intellegent systems (aiis'2017), Київ 2017, с. 101-108
13. Осауленко В.М. На шляху до створення інтелекту Philosophed 2015 міжнародна наукова конференція, ст. 47-49, Київ 2015.

Зміст

Вступ	14
1 Короткий огляд нейронауки та навчання в спайковій нейронній мережі	20
1.1 Короткий огляд стану нейронауки	21
1.2 Двомірна модель пластичності	38
1.2.1 Модель	39
1.2.2 Аналіз та результати	45
1.2.3 Обговорення	48
1.3 Проблеми спайкових нейронних мереж	50
1.4 Висновки	52
2 Просторова асоціативна пам'ять	54
2.1 Постановка задача та основні підходи	55
2.2 Асоціативна пам'ять на основі дендритних обчислень	63
2.3 Кодування зображення в розріджені представлення	83
2.4 Висновки	88
3 Часова асоціативна пам'ять	91
3.1 Просторово-часова активація біологічних нейронних мереж	91
3.2 Навчання послідовностей. Постановка задачі та основні підходи	98
3.3 Модель прогнозування послідовностей на основі сігма-пай нейрона та розрідженого представлення даних	103
3.4 Часова інтеграція послідовностей в нейрони зі збільшеною здатністю до активації	112
3.5 Висновки	121
4 Застосування моделей на прикладі навчання робота	123

4.1	Схема навчання робота Nao	124
Висновки		132
Додатки		134
1.	Список опублікованих праць за темою дисертації	134
2.	Довідка про впровадження в навчальний процес КПІ ім. Ігоря Сі- корського	137
3.	Акт впровадження в роботу лабораторії Logia м.Нансі, Франція . . .	138
Бібліографія		139

Вступ

Актуальність теми. За останні десятиліття кількість інформації значно збільшилися завдяки розвитку мережі інтернет та здешевленню засобів персональної генерації та поширення контенту. Тому значну увагу приділено алгоритмам, що обробляють та структурують великі масиви даних. Сюди включають алгоритми розпізнавання, кластеризації та, поки в зародковому стані, алгоритми зв'язування даних в причинно-наслідкову модель. Один із напрямків побудови та розвитку цих алгоритмів базується на дослідженні аналогічних біологічних моделей, так як люди легко справляються з такими задачами. Проте дані з нейронауки часто важко формалізувати в обчислювальні моделі, які не тільки будуть відтворювати спостережувані ефекти, але й пояснювати принципи обробки інформації.

Область нейронауки надзвичайно динамічно розвивається через нові технічні та теоретичні засоби дослідження, що виявляється в сотнях тисяч публікацій щороку. Основні напрямки, на які спирається дана робота, та які можуть потенційно породити нові, кращі алгоритми обробки даних включають роботу одного нейрона, математичні моделі якого були покращені такими вченими як E.Izhikevich, W.Gerstner, C.Koh; обчислювальні властивості дендритного дерева, досліджені в роботах B.Mel, P.Poirazi, G.Kastellakis, V.Sjostrom, T. Branco, M. London. В працях B.Olshausen, J.Hawkins, показано важливість розрідженої активації нейронної мережі, що узгоджується з теоретичними моделями асоціативної пам'яті в роботах D.Willshaw, A.Knoblanch.

Моделі асоціативної пам'яті популярні у 80-х роках, актуалізуються знову в роботах A.Graves та G.Wayne, через проблему «катастрофічного забування» в класичних нейронних мережах, що навчаються на основі оптимізації функції похибки. Нещодавно, моделі штучних нейронних мереж з пам'яттю набули особливого інтересу, так як потенційно можуть лягти в основу створення причинно-наслідкової моделі середовища для покращення розпізнавання та

прийняття рішень в задачах робототехніки.

Дослідження біологічних нейронних мереж є джерелом багатьох нових ідей і вказують на важливість представлення даних у вигляді розріджених багатомірних векторів для побудови штучних мереж з пам'яттю. Часто розглядається бінарна активація нейронів, зокрема в роботах Д.А.Рачковського, Р. Канєва. Операції над бінарними векторами, такі як додавання, перетин, асоціювання не тільки швидко виконуються при апаратній реалізації, але й відповідають біологічним експериментальним даним. Тому розробка моделей з таким представленням даних є перспективним напрямком як для кращого розуміння роботи реальних мереж, так і для створення більш ефективних штучних.

Отже, тематика обробки інформації та її збереження в розподіленому вигляді в нейронних мережах швидко розвивається, але існуючі алгоритми все ще далеко позаду природних біологічних систем. Тому, створення та дослідження нових моделей, які враховують останні біологічні дані можуть породити нові, кращі алгоритми та бути застосовані для задач розпізнавання та робототехніки.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконана у відділу прикладного нелінійного аналізу Інституту прикладного системного аналізу Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» відповідно до планів наступних науково-дослідних тем:

«Системи, моделі та рівняння з випередженням та запізненням. Дослідження математичних задач та застосувань» – номер державної реєстрації 0114U000205.

«Нові узагальнені математичні моделі з дискретним часом: клітинні автомати, нейромережі, дискретні рівняння та можливі їх застосування» – номер державної реєстрації 0119U002100.

Також робота виконувалася в лабораторії LORIA, м. Нансі, Франція, в рамках програми обміну Erasmus + та коледжі Каламазу, м.Каламазу, США.

Мета і задачі дослідження. Метою дослідження є розробка моделей

просторово-часової асоціативної пам'яті в моделях біологічних нейронних мереж; а також, проведення аналізу їх обчислювальних властивостей та можливості їх практичного застосування в задачах робототехніки.

Для досягнення мети дослідження поставлені й розв'язані такі задачі:

1. Побудовано двомірну модель навчання з різними часовими масштабами в спайковій нейронній мережі на основі сучасних даних з нейронауки.
2. Запропоновано нову модель асоціативної пам'яті з використанням моделі нейрона сігма-пай. Отримано аналітичні вирази ємності пам'яті. Порівняно з прийнятою на сьогодні моделлю асоціативної пам'яті Вілшова.
3. Застосовано нову модель асоціативної пам'яті для задачі прогнозування послідовностей, проведено теоретичний аналіз ємності пам'яті переходів та порівняно з подібною моделлю НТМ.
4. Побудовано модель представлення послідовності в бінарне розріджено-розподілене представлення на основі біологічних відомостей алокації (allocation) пам'яті.

Об'єктом дослідження є активація та навчання в моделях біологічних нейронних мереж.

Предмет дослідження – моделі асоціативної пам'яті у спайкових та сігма-пай нейронних мережах

Методи дослідження – як теоретичні (системний аналіз, узагальнення) так і практичні (проведення комп'ютерних симуляцій та обчислювальних експериментів). Технічна частина включає застосування теорії ймовірності та комбінаторного аналізу, теорії динамічних систем, теорії інформації та комп'ютерного моделювання.

Наукова новизна отриманих результатів.

Вперше:

- враховано дендритні обчислення для побудови моделі асоціативної пам'яті використовуючи модель нейрона сігма-пай, відрізняється тим, що має високу ємність збережених пар векторів;
- показано, що врахування здатності нейрона до активації приводить до часової інтеграції послідовності на масштабі часу секунд, що якісно відтворює біологічні ефекти такі як: збереження схожості, чутливість до порядку, доповнення послідовності та часову схожість.

Удосконалено:

- правила пластичності з різними часовими масштабами, відрізняється тим, що додано вплив повільної змінної сили зв'язку на швидку, що краще відповідає взаємодії молекулярних механізмів в синапсах;
- збереження контексту послідовності в моделях з розріджено розподіленим представленням даних, відрізняється тим, що замість колончастої організації використовується часова затримка активності, що спрощує архітектуру мережі.

Набула подальшого розвитку:

- теоретична оцінка ємності пам'яті на основі нейрона з дендритами, відрізняється тим, що досліджується не один нейрон, а мережа нейронів.

Практичне значення одержаних результатів полягає у побудові моделі асоціативної пам'яті високої ємності, яка може бути використана для швидкого пошуку по базі даних, а також у розпізнавання образів в робототехніці. Теоретичний аналіз показує важливість врахування дендритного дерева в моделі нейрона, що може бути використано практично в побудові нейроморфних архітектур. Часова інтеграція дозволяє представляти послідовність елементів (наприклад, візуальних рис) в окреме представлення, що може бути асоційоване та використане для розпізнавання послідовності. Розроблена схема, що дозволяє

навчати робота розпізнавати нові об'єкти в режимі онлайн. Хоча є ще багато невирішених проблем, проте, результати даної роботи мають практичне значення в побудові роботів з біологічно правдоподібними алгоритмами навчання.

За результатами стажуванні в лабораторії Loria, м. Нансі, Франція отримано акт впровадження результатів в роботу лабораторії. Також, впроваджено в навчальний процес КПІ ім. Ігоря Сікорського в курс «Нейронні мережі».

Особистий внесок здобувача. Всі наукові результати, що складають основний зміст роботи та становлять наукову новизну, отримані автором особисто. Більшість публікацій, результати яких увійшли в дану роботу автор опублікував особисто. Виняток становить робота [127] виконані у співробітництві з колегами, в якій автор побудував модель асоціативної пам'яті, провів комп'ютерні симуляції та отримав аналітичні вирази оцінки її ємності.

Апробація результатів дисертації. Наукові та практичні результати доповідались на семінарах та наукових конференціях:

- Осауленко В.М. На шляху до створення інтелекту Philosophed 2015 міжнародна наукова конференція, ст. 47-49, Київ 2015.

- Осауленко В.М. Тестування простих моделей з дендритами для розрідженого бінарного представлення зображень. XVII Міжнародна науково-технічна конференція Штучний інтелект та інтелектуальні системи. Artificial intelligence and intellegent systems (aiis'2017), Київ 2017, с. 101-108

- Osaulenko V.M. New ideas to test for artificial neural networks inspired from neuroscience Міжнародна наукова конференція «Сучасна інформатика: проблеми, досягнення та перспективи розвитку» Інституту кібернетики імені В.М.Глушкова, 13-15 грудня 2017, с.82-83

- Osaulenko V.M. Learning of higher-order correlations improves capacity of an association memory. Nonlinear analysis and applications 4rd International Conference of memory of corresponding member of National Academy of Sciences of Ukraine Valeriy Sergeevich MELNIK (1952 - 2007), NTUU KPI, Kyiv, Ukraine, April 04-06, 2018, p.76

- Osaulenko V.M. Model for sequence prediction based on dendritic spatiotemporal integration. 20-th International conference on System Analysis and Information Technology SAIT 2018, May 21-23, 2018 pp.36-37 Institute for Applied System Analysis of National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute Kyiv, Ukraine

- Osaulenko V.M. Dendritic nonlinearity provides high capacity sequence memory FENS FORUM of neuroscience 2018, Berlin, p.368

- Makarenko O.S., Osaulenko V.M. Application of cellular automates in some models of artificial intelligence. Конференція IEEE SAIC-2018, 2018, 1-4. 10.1109/-SAIC.2018.8516837

- Науковому семінарі при Навчально-науковому комплексі “Інститут прикладного системного аналізу” КПІ ім. Ігоря Сікорського (Київ, 2018).

Публікації. Основні результати дисертаційної роботи опубліковано в 13 наукових працях у тому числі 6 статей у наукових фахових виданнях (з них 2 статті у виданнях іноземних держав, 4 статті у виданнях України, які включені до міжнародних наукометричних баз), 7 тез доповідей в збірниках матеріалів конференцій.

Структура та обсяг дисертації. Дисертація складається зі вступу, чотирьох основних розділів, висновків, списку використаних джерел та додатків. Робота викладена на 158 сторінках і містить 132 сторінок основної частини, 50 рисунків, 1 таблицю і список використаних джерел із 177 посилань.

Розділ 1

Короткий огляд нейронауки та навчання в спайковій нейронній мережі

Нервова система забезпечує багатоклітинний організм здатністю до виживання. Це відбувається завдяки можливості трансформувати різні типи енергій (механічну, електромагнітну, гравітаційну) в електричні імпульси нейронів, що інтегруються з мільйонів сенсорів в цілісну модель середовища. Модель створює передбачення майбутнього, що дозволяє прийняти правильні рішення та підготувати відповідні моторні команди. В основі роботи моделі лежить розпізнавання образів на основі довго- та короткотривалої пам'яті. Також, важлива інформація виділяється через сигнали підкріплення, такі як насолода, біль, новизна, подив та інші, що викликають емоційну реакцію. Залишається ще не вирішеною задача, як така модель навчається в режимі реального часу, має велику ємність та хорошу здатність до розпізнавання. Штучна система з такими ж можливостями, знайшла б повсюдне застосування в людській діяльності, і тому триває пошук шляхів її створення. Один з них це безпосереднє дослідження роботи мозку людини, збирання всіх необхідних біологічних фактів та побудова всеохоплювальної теорії мозку, з урахуванням теорії інформації та фізичних принципів. Інший шлях алгоритмічний, орієнтований на вирішення задач, таких як класифікація чи робота з текстом. Важливо поєднувати ці два шляхи, біологічний та математичний, щоб зрозуміти яким чином інформація кодується і обробляється в живих організмах для створення розумних машин.

1.1 Короткий огляд стану нейронауки

З початком комп'ютерної ери 40-х років 20 століття, люди сподівалися, що на-решті буде відтворений розум в іншому носії, окрім біологічного, проте задача виявилася надзвичайно складною. Питання, яким чином людина здатна пізна-вати, мислити, усвідомлювати вивчали філософи протягом багатьох століть. По сьогоднішній день досі немає всіма погодженого визначення інтелекту та шляхів його створення, але схоже, що теперішнє століття дасть довгоочікувані відповіді [126].

Найкращий спосіб зрозуміти щось це його створити. Ще в часи античності люди розмірковували про штучні створіння, які інтелектуально зрівнюються з людьми. В часи середньовіччя намагалися описати роботу людини, як дуже складні механізми. Але по справжньому почали науково вивчати нервову си-стему з кінця 19 століття, коли Гольджі винайшов метод забарвлення нейронів, які в звичайних умовах є прозорими. Використавши цей метод, іспанський вче-ний Сантього Кахаль описав перші принципи організації нервової тканини, чим самим поклав початок нейронауки. Нервова тканина складається з клітин, які назвали нейронами. Вони мають довгі відростки: аксони та дендрити, через які нейрони з'єднуються між собою в мережі та обмінюються інформацією завдяки електричним імпульсам. Така взаємодія мільйонів нейронів забезпечує сприйня-ття, пам'ять та інтелектуальну поведінку. Важлива особливість, на яку рідко звертають увагу, що нейрони живуть протягом всього життя організму. Це дає неперервність пам'яті та особистості, та впливає на пластичність зв'язків, тобто яким чином нейрони створюють, видаляють та підтримують зв'язки між собою.

Одна із перших спроб формально описати роботу нейрона була зроблена в 1907р. Нейрон був описаний як інтегратор, і якщо його напруга перевищує поріг, то він генерує електричний імпульс і перезаряджається до початкового стану [1].

$$\tau \frac{dV(t)}{dt} = I(t) \text{ Якщо } V(t^f) > V_{thr} : s(t^f) = 1, V(t^f) = V_0 \quad (1.1)$$

Тут $V(t)$ напруга нейрона, яка змінюється з часом під дією вхідного сигналу $I(t)$ (тут і далі розмірність вольт). Якщо напруга більше певного порогу V_{thr} , то нейрон активується $s(t^f) = 1$ і напруга скидається до початкового стану. Така модель включає здатність нейрона сумувати вхідний сигнал від багатьох нейронів і генерувати імпульс за певної умови. Проте, вона не включає згасаючий доданок, що повертає нейрон в стан рівноваги за відсутності вхідного стимулу як це спостерігається в реальних нейронах.

Наступна спроба була зроблена В. Маккалохом та В. Піттсом в 1943 [112]. Математично, нейрон рахує суму вхідних зв'язків x_i помножену на їх силу w_i і порівнює з порогом b .

$$y = f(\sum w_i x_i), \quad f(z) = \begin{cases} 1, z \geq b \\ 0, \text{otherwise} \end{cases} \quad (1.2)$$

Ідея схожа з попередньою моделлю, нейрон як функція від входу, порівнює свій стан з граничним значенням і приймає два значення: 0 та 1. Проте, тут явно задані сили зв'язку та немає пам'яті відносно попередніх станів. Схематично модель зображена на рис.1.1. Саме з цієї моделі почався розвиток нейронних мереж і після конференції 1956 року виник термін "штучний інтелект".

Далі, Розанблант в 1958 році заклав модель Маккалоха — Піттса в основу алгоритму персептрона, і показав, що вона здатна робити лінійну класифікацію через зміну ваг w_i . Проте, в 1969 році Мінські та Пеперт опублікували роботу [117], в якій показали обмеженість алгоритму, зокрема, неможливість нелінійної класифікації. Тоді було зрозуміло, що для більш складних задач потрібно обчислювати інформацію в багатьох шарах нейронів, проте виникала проблема з навчанням таких мереж. В даному напрямку працювали і українські вчені, зокрема, О. Івахненко та О. Різник, що досліджували навчання порогу актива-

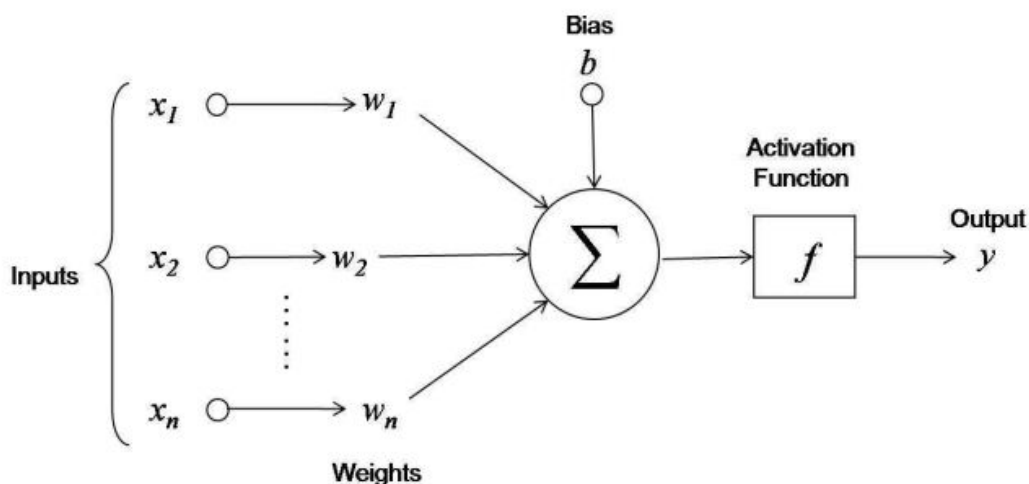


Рис. 1.1: Схематичне зображення нейрона Маккалоха — Піттса

ції нейронів. Тим не менш, розвиток пішов іншим шляхом. Відроу і Хофф [171] у 1960 розширили модель нейрона, так, щоб вихід був реальним значенням і рахувався через сігмоїду $\sigma(x) = \frac{1}{1+\exp(-x)}$. Пізніше, в кінці 60-х та в середині 70-х був винайдений спосіб навчання параметрів багат шарових мереж через алгоритм зворотного поширення похибки (backpropagation). Більшого поширення він здобув після його перевинайдення в роботі Румелхарда, Хінтона, Вілсона [150] 1986 року. З тих пір почався новий виток розвитку нейронних мереж, зокрема, напрямку навчання з учителем. Також, продовжувався обмін ідей між нейронаукою та формальними моделями. На основі робіт 60-х років Хьюбеля і Візеля [72, 73] дещо пізніше у 1980 році була створена нейронна мережа під назвою неокогнітрон [49], яка лягла в основу всіх подальших моделей згорткових мереж.

В період кінця 80-х до початку 10-х років нейронні мережі знайшли застосування в розпізнаванні зображень, звуку, в задачах обробки природної мови та виникло багато їх варіацій. Обмежені машини Больцмана, автоенкодера, згорткові нейронні мережі доповнювали і значно покращували стандартні архітектури. Проте, поруч з нейронними мережами виникли інші алгоритми, під конкретні задачі, які часто працювали краще. Наприклад, метод опорних векторів (SVM) для класифікації, або приховані марківські моделі (НММ) для задач

з часовими послідовностями довгий час показували найкращу ефективність. Однак, нещодавно стався новий виток розвитку нейронних мереж. В 2012 році в роботі [95] показали, що багатошарова згорткова нейронна мережа навчена на великій кількості зображень, зміненою активаційною функцією для нейрона та використанням нового методу регуляризації покращує похибку класифікації з 26% до 15.3%. Такий великий стрибок у точності, викликав великий інтерес у спільноти і з тих пір виникло багато варіацій даної мережі, які досягають ще кращих результатів. Сукупність методів використання багатошарових нейронних мереж з використанням великої кількості даних називають DeepLearning.

Незважаючи на великий успіх в прикладних задачах, більшість дослідників погоджуються, що одних методів навчання з учителем Deep Learning замало в довгостроковій перспективі для створення розумних машин. Відмічається, що ключовим напрямком має стати навчання без учителя, коли нейронні мережі використовуючи принципи самоорганізації, інтегрують та перетворюють інформацію таким чином, щоб відтворювати статистичні закономірності та розподіл ймовірностей вхідних даних. Гіпотетично, це має перетворити дані в зручний спосіб, і тоді сигнал підкріплення з середовища створює необхідні асоціації. Тому поруч з класичними нейронними мережами розвиваються інші підходи, з іншими ідеями навчання [109].

Цікаво, що фізичні закони можуть бути еквівалентно описані як причинно-наслідкові рівняння, або як оптимізація певної величини. Наприклад, в механіці можна застосувати рівняння руху або оптимізувати функцію Лагранджа. Схоже, що такий самий принцип може бути використаний і в області нейронних мереж. Класичні нейронні мережі слідує напрямком оптимізації, зменшуючи функцію похибки. Альтернативні підходи, зокрема, з обчислювальної нейронауки, будують висхідні причинно-наслідкові моделі, які легше аналізувати, але важче навчати.

В 90-х роках почали активно розвиватися спайкові нейронні мережі, які ще називають третім поколінням нейронних мереж [107]. Перше покоління викори-

стовувало нейрон Маккалоха — Піттса, друге – лежить в основі Deep Learning, третє – використовує складніші моделі нейронів, наближені до реальних, що керують станом нейрона в часі в залежності від вхідного збудження та внутрішньої динаміки. Дані нейрони мають бінарний вихід, як і в першому поколінні, але сам стан нейрона може описуватися багатьма іншими змінними, що відслідковують пам'ять минулої активації.

Найпростішою спайковою моделлю нейрона є модель дисипативна інтегруй-активуй (leaky integrate and fire), яка описується наступним чином:

$$\begin{aligned}\tau \frac{dV(t)}{dt} &= -V(t) + I(t) \\ V(t^f) > V_{thr} : s(t^f) &= 1, V(t^f) = V_0\end{aligned}\tag{1.3}$$

де $V(t)$ напруга на мембрані, що описує стан нейрона, $I(t)$ зовнішнє збурення, або від електрода, або від інших нейронів, V_{thr}, V_0 - параметри порогу та початкового стану. Якщо $V > V_{thr}$ нейрон стає активний $s = 1$, або ще кажуть, генерує спайк. Модель якісно враховує суть роботи біологічного нейрона, але не має пам'яті про історію активності, так як кожний спайк скидає напругу до початкового значення і стирає пам'ять. І тому, ця модель не може відтворити ряд спостережуваних феноменів, таких як адаптація чи пучкоподібна активність. Інші моделі використовують дві змінні для опису динаміки нейрона і є значно точнішими. Типовими прикладами є моделі Іжікевича [79] та адаптивна експоненційна інтегруй-активуй [24]. Остання записується наступним чином:

$$\begin{aligned}\tau \frac{dV}{dt} &= f(V) - w + I \\ \tau_w \frac{dw}{dt} &= a(V - b) - w \\ V(t^f) > V_{thr} : s(t^f) &= 1, V(t^f) = V_0\end{aligned}\tag{1.4}$$

Де $f(V) = g(V - E) + g \exp \frac{V - V_T}{\Delta_T}$, та $a, b, g, E, V_T, \Delta_T, V_{thr}$ - параметри. Така модель є дещо складнішою, за просту одномірну, але є значно простішою за детальну модель Ходжкіна-Хакслі - першу модель, яка математично описала

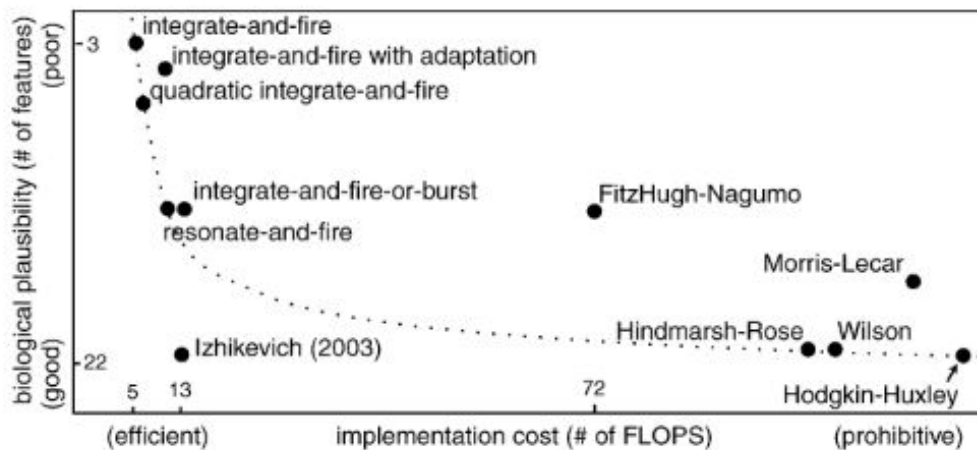


Рис. 1.2: Моделі активації спайкових нейронів. Залежність біологічної подібності поведінки моделі (кількість рис) від її обчислювальної складності (FLOPS). Модель Ходжкіна—Хакслі внизу справа, найбільш детальна, але й найбільш обчислювально затратна. [77]

роботу нейрона, і яку було відзначено Нобелівською премією. Різні моделі та їх ефективність добре ілюструє рис. 1.2 з роботи Іжікевича [77]. По осі абсцис відкладено точність відтворення біологічних феноменів, по осі ординат - обчислювальну складність. Як видно модель Ходжкіна-Хакслі та інтегруй-активуй знаходяться по різних кінцях, з багатьма моделями посередині.

Математичний апарат динаміки моделі нейронна розроблений досить досконало, який дає змогу відтворити активацію біологічного нейрона. Проте, реальний нейрон надзвичайно пластичний, тобто змінює свої параметри у відповідь на зовнішнє середовище. В ці параметри включаються зв'язки від інших нейронів та змінні, що керують внутрішньою динамікою. Для повного опису роботи біологічного нейрона потрібно знати просторово-часовий розподіл та тип іонних каналів на поверхні мембрани, на кінцях аксонів та на дендритних шипиках, а також, скоріш за все, молекулярно-генетичні процеси, які керують динамікою перерозподілу цих каналів. В ідеальному випадку, поки що, такий опис не є можливий. Також, всі правила пластичності не відомі, і тому на рівні нейронних мереж, коли нейрони впливають друг на друга, не вдається точно описати динаміку, яка б повністю відтворювала біологічну. В рамках Human Brain Project

[111] було створено реконструкцію $0,3\text{мм}^2$ кори головного мозку, в якій описали тип та розташування близько 30 тисяч нейронів, з їх зв'язками, електричними та морфологічними властивостями. Однак, потрібно додавати феноменологічні правила для зміни параметрів цих нейронів з часом.

Окрім самої моделі нейрона, важливою є його морфологія, тобто форма дендритного дерева. За останні 20 років, отримано вичерпні докази, що дендритні сегменти не лише збирають зовнішні сигнали, але й виконують власні обчислювання. Було показано, що завдяки наявності потенціал або кальцій залежних іонних каналів, інтеграція вхідної активності відбувається нелінійно, тобто сума індивідуальних ваг синапсів не дорівнює сумарній активації, яку вони створюють. Більш детально дендритним обчислюванням присвячено другий розділ, де показано в деталях біологічні факти, та як вони приводять до вищої ємності пам'яті в мережі.

Принципи кодування інформації

Напрямок обчислювальної нейронауки, що використовує спайкові нейронні мережі стикається не тільки з проблемами навчання мережі, але і з більш глобальними, наприклад, як кодується інформація і яку задачу повинні виконувати нейрони.

Якщо вставити один електрод в нейрон, а інший тримати ззовні, то між ними виникає напруга, яка час від часу змінюється стрибками. На осцилографі ці стрибки нагадують шипи, що з англійської перекладається як спайк (spike), тому часто замість електричного імпульсу чи потенціалу дії говорять, що нейрон генерує спайк. У відповідь на певні стимули частота спайків росте, тому перше природне припущення було, що нейрон кодує інформацію частотою спайків. Особливо це спостерігається в периферійній нервовій системі, що відповідає за сприйняття сенсорної інформації та реалізацію моторних команд. Чим більша частота активації нейрона, тим сильніший стимул, або тим сильніше, напри-

клад, стискається м'яз. Також, Хюбель і Візель [71] показали, що нейрони у візуальній корі головного мозку активуються з більшою частотою на свій характерний стимул, і частота зменшується, якщо стимул змінюється. На рис. 1.3 показано криву залежності частоти спайків від орієнтації нахиленої лінії, що потрапляє в рецептивне поле нейрона у первинній візуальній корі головного мозку V1. Як видно, при зміні нахилу лінії, частота активації нейрона падає.

З часом були отримані нові результати кодування інформації в реальних нейронних мережах, що показали, що надзвичайно важливо не тільки частота активації нейрона, але й час індивідуальних спайків. Зокрема, форма пластичності STDP, про яку йтиметься далі, залежить від часу активації двох нейронів. Також, серйозним аргументом проти частотного кодування в сенсорній корі головного мозку є спостереження, що людина розпізнає зображення близько за 100 мс, при тому, що інформація поширюється через 6-8 шарів нейронної мережі. Враховуючи час передачі спайку по аксонах (3-10 мс), генерацію спайку (2-5 мс), та передачі сигналу в синапсах (1-2 мс), для узгодження зі швидкістю обробки інформації один нейрон може передати максимум 1-2 спайки з яких неможливо визначити частоту. Тому, сходяться на думці, що питання взагалі має бути поставлено по іншому, зокрема, а що ж кодує нейрон?

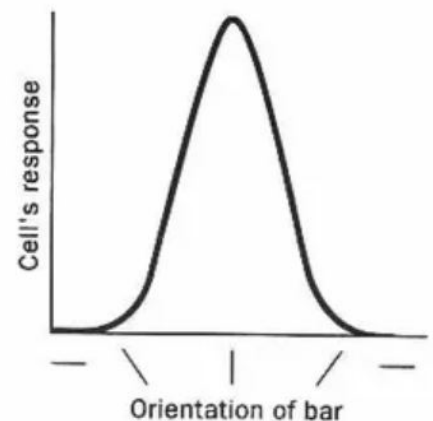


Рис. 1.3: Крива частоти активації нейрона в залежності від стимулу

В 90-х роках з'явилася концепція "нейрона бабусі" суть якої в тому, що якщо людина дивиться на свою бабусю, то певний нейрон в мозку активується який її закодує. Така ідея пов'язана з експериментальним спостереженням, що, дійсно, міряючи активність певного нейрона, спостерігається велика кореляція зі стимулом. Таке кодування називається локальне, так як стимул кодується одним нейроном. Проте, згідно сучасного уявлення, інформація кодується групою ней-

ронів, що називають популяційним кодуванням. В експерименті з мавпами зчитувалася активність моторної кори головного мозку при русі руки [52], і було показано, що, наприклад, при русі вгору, певна група нейронів активна, а при русі вниз, активна інша група. Один нейрон з групи може активуватися і для різних напрямків, але вся сукупність нейронів унікально кодує рух.

Пізніше, було показано, що принцип популяційного кодування може відбуватися не тільки в моторній корі, але і в сенсорній. Спостерігаючи за кодуванням в корі мозку, що відповідає за дотик у мишей, було показано, що нейронна мережа у відповідь на стимул активується групами нейронів, і що ці групи розподілені по всій мережі, і що активність сильно розріджена, тобто відношення кількості активних нейронів до всіх нейронів значно менше одиниці [105]. Такий принцип називають розрідженим розподіленим кодуванням, і воно присутнє в інших частинах кори головного мозку у різних тварин [125]. На рис. 1.4а червоним зображено активні нейрони в різні моменти часу, ілюстрація відтворює дані з активності соматосенсорної кори гризунів. Цікаво, що в цьому випадку біологічні спостереження узгоджується з теоретичними результатами теорії інформації. Моделі, що використовують розріджене кодування збільшують ємність асоціативної пам'яті, а також мають ряд інших обчислювальних переваг [68, 82, 83]. Про це буде йти в другому розділі.

Також, важливо як розподілена в часі розріджена активність. Нейронні мережі працюють ритмічно, тобто активність розподілена не хаотично, а згрупована в моменти часу, в які синхронно багато нейронів активні. Про таку ритмічність знали вже давно, ще внаслідок вимірів електроенцефалограми, проте з появою мультиелектродних зчитувачів, які міряють активність безпосередньо з великої кількості нейронів, порядку 100, змогли детальніше дослідити часовий розподіл генерації спайків. В результаті виникла гіпотеза про пакетну передачу інформації [106], де інформація з однієї області в іншу передається пакетом спайків. На рис. 1.4б показано активацію 40-ка нейронів (по вертикальній осі), протягом двох секунд вимірювання. Видно, що активність згрупована в певні

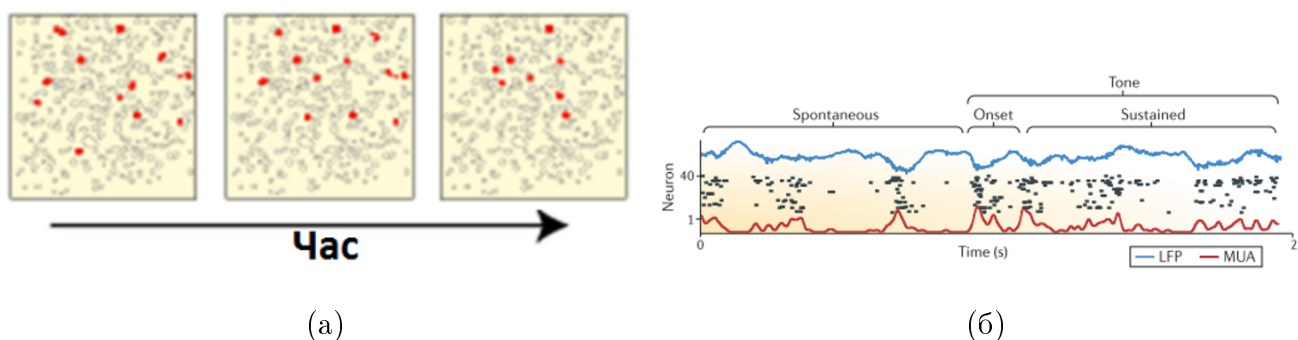


Рис. 1.4: а) Ілюстрація розрідженої активності нейронів. В кожен момент часу активується лише мала частина нейронів. Дані з соматосенсорної кори гризунів [105]. б) Активація 40-ка нейронів (по вертикальній осі), протягом двох секунд вимірювання. Червоним показано усереднену активацію популяції нейронів. Спостерігається групова активність нейронів при реакції на стимул, так і при спонтанній активації [106].

моменти часу. Синхронна активність має вирішальну роль в навчанні, увазі та свідомому сприйнятті [160].

Динаміка активності реальних мереж надзвичайно складна, і до кінця не зрозуміла. Вважається, що зміна активності в часі, відображає динаміку реального світу, зокрема, виконується фундаментальна функція - навчання послідовностей. Можна говорити про часову асоціативну пам'ять, яка з'єднує різні популяції нейронів, що кодують події в різні моменти часу. Наприклад, м'яч котиться і падає з столу. Нейрони, які кодують м'яч у різні моменти часу, з'єднуються в послідовному порядку. І в майбутньому, коли м'яч тільки почне котитися, нейрони активують всю послідовність раніше, ніж реальність, і буде відбуватися передбачення, що м'яч має впасти. Вважається, що саме завдяки навчанню послідовностей, що відображають причинно-наслідкові зв'язки, з'являється базове уявлення про фізичні закони [62], як то рух по прямій чи рух сонця зі сходу на захід.

Як доказ, багато експериментальних даних свідчить про те, що послідовна активність нейронної мережі відтворює послідовність сенсорного досвіду. Наприклад, у гіпокампі мишей перед входом у відомий лабіринт, відбувається послідовна активація нейронів, що були колись активні при переміщенні в ла-

біринті [15]. Таким чином, відбувається елементарне передбачення через прокручування досвіду. Нейрони у звуковій корі, також активуються послідовно, відтворюючи послідовність звуків, що вони кодують [87]. Третій розділ роботи присвячений навчанню послідовностей, тому більше біологічних деталей буде представлено згодом, зокрема, які можливі біологічні основи передбачення та інтеграції послідовності.

Потік інформації в біологічних нейронних мережах

Важливо зазначати більш глобальні принципи перетворення інформації в нервовій системі. Близько 1 млрд рецепторів з очей, шкіри, м'язів, внутрішніх органів, вух, нюху, смаку, вестибулярної системи сприймають інформацію з зовнішнього середовища і перенаправляють в подальшу обробку. Мета первинних центрів нервової системи інтегрувати всю цю інформацію в єдину модель, що відображає регулярності в середовищі. Такі регулярності можуть бути просторовими, наприклад, у візуальній інформації відображається наявність нахилених граней об'єктів, або часовими, наприклад, причинно-наслідкові події, як рух по прямій.

Основні етапи поширення інформації зазначено на рис. 1.5 з роботи Дж. Лісмана [103]. Спершу, інформація проходить первинну обробку в сенсорних органах. Наприклад, у сітківці відбувається зменшення розмірності з 100 млн фоторецепторів до 1 млн гангліальних клітин. Далі, інформація направляється через таламус, який діє як передавач і модулятор в кору головного мозку. А також по паралельним шляхам, передаються сигнали в інші підкортикальні структури для швидкої несвідомої обробки. Такі короткі шляхи дають швидку реакцію на загрозові стимули, і тому організм має більше шансів вижити.

Кора головного мозку виділяє регулярності у вхідних даних та співвідносить їх між собою. Гіпокамп, завдяки наявності багатьох рекурентних зв'язків створює асоціації між високорівневими концепціями, що відіграє важливу роль у

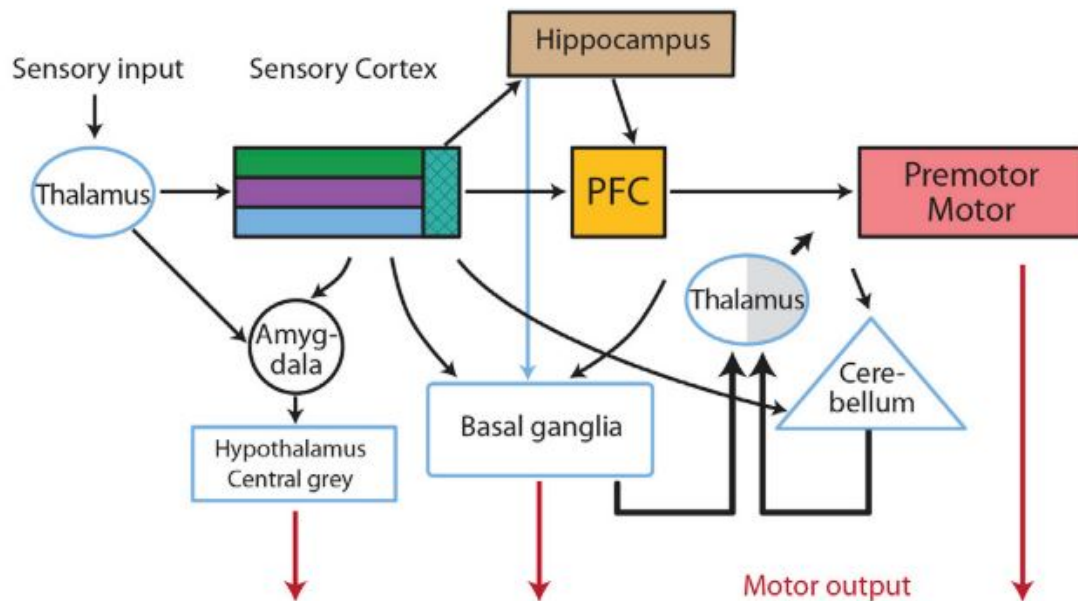
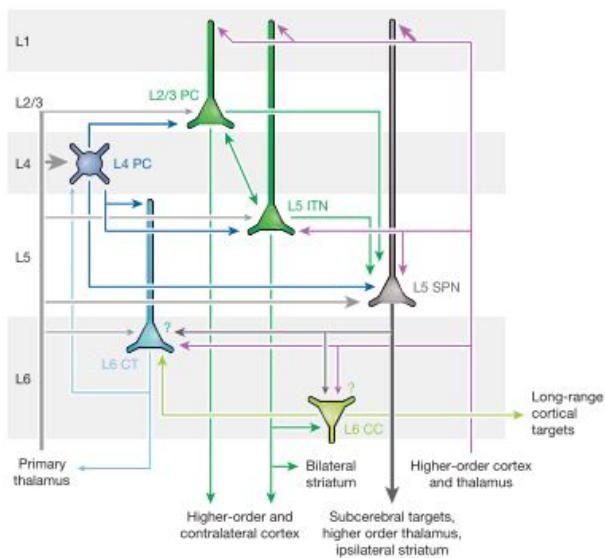


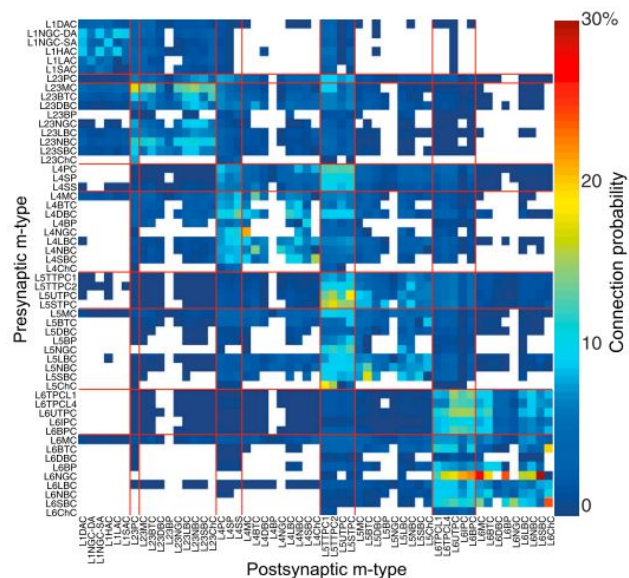
Рис. 1.5: Поширення інформації в основних областях мозку [103].

формуванні епізодичної пам'яті. Префронтальна кора є ще вищим ієрархічним рівнем, який поєднує та зберігає інформацію в робочій пам'яті з різних частин мозку. Префронтальна кора найменш вивчена, але відомо що вона приймає важливу роль в прийнятті рішень, формує довгострокові передбачення та зберігає типову поведінку, в контексті багатьох ситуацій, що формує особистість. Вся ця обробка інформації слугує для виконання правильних моторних команд, тобто керує поведінкою тварини. Моторна кора разом з базальними гангліями та мозочком визначають, яка команда і як саме буде виконана.

Підкортикальні структури вже досить добре вивчено. Отримано інформацію про їх зв'язки та функції. Але, досі немає розуміння про принципи обробки інформації в корі головного мозку. Нейрони в корі організовані латерально у шари, яких виділяють від 5 до 9 залежно від області мозку, та у колонки, вертикально організованих нейронів. На рис. 1.7 схематично показано організацію нейронів в кортикальну колонку висотою 1.6 мм та діаметром 23 мкм. Червоним показано дендрити, які простягаються вздовж всіх шарів до поверхні кори, а синім – аксони, які виходять до інших областей. Всього в такій колонці близько 100 нейронів. Залишається відкритим питанням чи така організація має фун-



(а)



(б)

Рис. 1.6: а) Зв'язки у кортикальних шарах [60], б) Ймовірність зв'язків між шарами [111]

кціональну роль, чи це просто наслідок процесів розвитку нервової системи. В. Маункастл показав [119], що нейрони в межах однієї кортикальної колонки мають схоже рецептивне поле, тобто реагують на один і той самий стимул з певної області рецепторів схожим чином. Було висунуто припущення Ф.Кріком, що реагуючи схожим чином і проектуючи свої аксони в різні частини мозку, нейрони в межах кортикальної колонки можуть передавати одну і ту ж інформацію багатьом іншим нейронам.

Виділяють канонічні нейронні зв'язки локально в межах шарів [35, 60], які повторюються по всій корі 1.6а. Таламус проектує аксони в основному в шари 4 та 6, у відношенні близько 1 до 100, тобто на 1 нейрон з таламусу, припадає 100 нейронів у корі. Припускається, що завдяки більшій кількості нейронів в корі забезпечується розріджена активація [125]. Часто виділяють основні напрямки передачі активності. Наприклад, нейрони з 4-го шару передають сигнал в шар 2/3 який потім передає сигнал в шари 4 та 2/3 іншої колонки, та проектує зв'язки в шар 5 своєї колонки. Шар 5, далі посилає сигнал на ядра таламус вищого порядку, або прямо на ієрархічно вищі рівні кори. Однак, останнім часом

приходять до згоди, що такі схематичні зображення зв'язків дуже спрощений погляд на кору, опис насправді ймовірнісний як видно з рис. 1.66. Нейрони з одного шару роблять зв'язки майже зі всіма іншими шарами [111] з різною ймовірністю. Проте, залишається специфічність, коли певні типи нейронів зовсім не створюють зв'язки з іншими типами нейронів. Також, важливо не тільки чи два нейрони зв'язані, а де саме на поверхні нейрона утворюються зв'язки. Наприклад, один тип нейронів інгібіторів створює контакти на самій клітині або на виході аксону, що може повністю скасувати активність. Інший тип інгібіторів, створює зв'язки на дендритах, тим самим лише модулюючи активність нейрона.

Така складність та різноманітність породжує багато можливих теорій та гіпотез, які намагаються пояснити суть обчислень в корі. Одна з гіпотез, що нейрони в первинних сенсорних областях вивчають розподіл ймовірностей вхідних даних. Нейрони спеціалізуються на певних стимулах, наприклад, похилі лінії, кути, решітки, які зустрічаються найчастіше у візуальній інформації. Причому, кодування відбувається з втратами, подібно до алгоритмів квантизації векторів (vector quantization). Особливої уваги заслуговують алгоритми розрідженого кодування (sparse coding) [43] та алгоритм незалежного аналізу компонент (ICA) [12]. Перший будує набір словників-зображень, лінійна комбінація малої кількості яких, добре відтворює зображення. Другий будує набір фільтрів, які генерують зображення з максимальною взаємною інформацією до вхідного. І перший і другий алгоритми породжують фільтри селективні до ліній різного нахилу, кутів, подібно до спостережуваних в візуальній системі мишей та мавп.

Характерною рисою поширення інформації в корі є ієрархічна організація нейронних мереж [42]. Верхні рівні ієрархії інтегрують інформацію з нижніх і мають доступ до більших рецептивних полів, а отже об'єднують сигнали з більшої кількості рецепторів. Також, важливим є факт наявності великої кількості зворотних зв'язків, інколи навіть більшої, ніж прямих. Точно невідомо чому нейрони з вищих рівнів надсилають багато аксонів до нижніх, проте при-

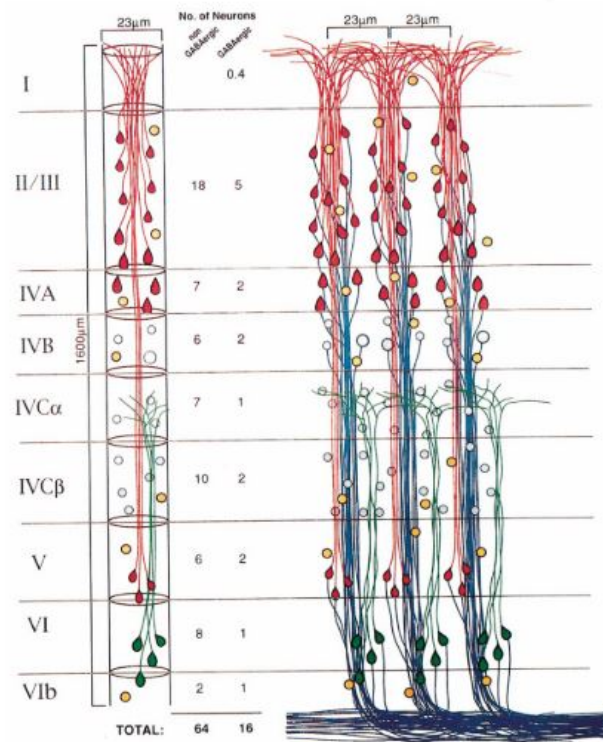


Рис. 1.7: Кортикальна колонка [119]

пускають багато гіпотез, таких як необхідність робити передбачення [145, 48], уточнення апріорного розподілу в рамках баєсівського підходу [100, 86], або ж передача контексту для покращення розпізнавання чи передбачення послідовностей [161, 149].

Загальна картина наступна, нейронна мережа, влаштована модульно та ієрархічно. Кожний рівень обробляє інформацію, та обмінюється з іншими рівнями. При цьому часто використовується розріджено розподілений код який передається пакетами. В рамках одного рівня відбувається навчання без учителя коли відображається вхідний розподіл ймовірностей. Різні рівні з'єднуються асоціативними механізмами подібними до хеббівського навчання.

Дана робота слідує поглядам, що існують абстрактні принципи обробки інформації, які через еволюцію знайшли біологічну реалізацію. І тому, основна мета дослідження мозку, це виділення цих принципів, та спроба їх втілення у іншому матеріальному носії.

Нейронна пластичність

Ключовим питанням в дослідженні мозку є питання навчання, або в термінах нейронауки, питання нейронної пластичності. Вже відкрито багато молекулярних механізмів, яким чином синапс одного нейрона з'єднується, підсилюється, або видаляється від дендрита іншого нейрона. Для спрощення опису зміни сили зв'язку, користуються феноменологічними правилами, які якісно та кількісно описують навчання, не вдаючись до біологічної реалізації. Найбільш відомим феноменологічним правилом є правило Хебба [67], яке формулюється так: якщо нейрон А бере участь в активації нейрона В, то він підсилює зв'язки до нього. Математично це може бути виражено як $\Delta w_{ij} = x_i y_j$, де x, y активації двох нейронів, а Δw зміна сили зв'язку між ними. Пізніше виникло багато модифікацій цього правила, які здебільшого намагаються усунути два найсуттєвіших недоліки: а) правило Хебба працює як позитивний зворотній зв'язок, тому два нейрони будуть підсилювати зв'язок між собою до нескінченності, б) правило не передбачає механізму зменшення сил зв'язків за умови додатних активацій.

Найвідоміший продовженням правила Хебба стало правило ВСМ [16], яке зменшує вплив недоліків та краще співвідноситься з експериментами. Основне нововведення це додавання гомеостатичної пластичності, яке враховує активацію нейрона на довгих проміжках часу, та повертає нейрон в оптимальний режим роботи.

$$\begin{aligned}\Delta w &= xy(y - \theta) \\ \tau \frac{d\theta}{dt} &= -\theta + y\end{aligned}\tag{1.5}$$

Тут з'являється доданок який враховує xy^2 взаємодію, та θ що відслідковує усереднену активність нейрона y . Іншим, також відомим правилом, що уникає насичення зв'язків, є правило Оджа [124], яке обчислювально знаходить перші принципові компоненти у вхідних даних.

Проблема цих правил навчання в тому, що вони не враховують час надхо-

дження спайків, від якого може залежати сила зв'язку. В 1997 році показали [110], що зв'язки між нейронами А та В, змінюються в залежності від відношення часу надходження спайку від нейрона А до часу активації нейрона В. Зв'язок підсилюється, якщо активація від нейрон А вплинула на нейрон В і він після цього став активний, але якщо нейрон В став активний до того, як надійшла активація від нейрона А, то зв'язок послаблюється. Величина зміни зв'язку залежить від відносного часу активації двох нейронів: чим ближча в часі активація, тим більша зміна. Це можна виразити правилом спайк-часово-залежної пластичності (STDP):

$$\begin{aligned}\frac{dw_{ij}}{dt} &= A_+x_i(t)s_j(t) + A_-y_j(t)s_i(t) \\ \frac{dx_i}{dt} &= -\frac{x_i}{\tau_x} + s_i \\ \frac{dy_j}{dt} &= -\frac{y_j}{\tau_y} + s_j\end{aligned}\tag{1.6}$$

де A_+, A_-, τ_x, τ_y - константи, $s_i = \sum_{i=1}^n \delta(t - t_i^f)$ - сума дельта функцій, що враховує всі спайки нейрона. Дане правило краще передбачає результати експериментів, але пізніше було показано [156], що правила пластичності значно складніші, вони мають враховувати і відносний час між спайками, і частоту спайків, і місце на дендритному дереві, і активацію інших сусідніх чи віддалених синапсів. Реальні механізми навчання біологічних нейронних мереж надзвичайно різноманітні і складні, тому дуже важко правильно їх всіх врахувати в одну модель.

Триплетна STDP частково виправляє зазначені проблеми, і в поєднанні з гетеросинаптичними гомеостатичними механізмами може підтримувати стабільну активацію мережі, яка може працювати як автоасоціативна пам'ять [177]. Час навчання та ємність мережі показують значно гірші показники порівняно з іншими, менш біологічно деталізованими, моделями асоціативної пам'яті. Проте, як писалося вище, важливо виділити, які обчислювання виконують нейронні мережі, щоб легше було зрозуміти, які правила навчання мають бути. Вважає-

ться, що ключовим механізмом пам'яті є створення асоціацій, тому в наступному підрозділі розглядається біологічно правдоподібна модель, що поєднує різні види навчання нейрона для зв'язування нейрона з активною популяцією.

1.2 Двомірна модель пластичності

Спайкові нейронні мережі вважаються перспективними в побудові адаптивних систем з навчанням і пам'яттю, що відтворюють характеристики живих організмів. Використання математичних моделей, що відтворюють динаміку реального нейрона призводить до складної і різноманітної поведінки в штучних системах. Головна задача – це вияснити, як з хаотичної динаміки з'єднаних в мережу тисяч нейронів отримати здатність запам'ятовувати велику кількість стимулів і потім їх відтворювати і розпізнавати. Ми ще знаходимося на ранньому етапі, так як досі залишається до кінця не зрозумілим яким чином кодується інформація в таких нейронних мережах і в мозку в цілому.

Реальний біологічний нейрон використовує багато механізмів, які забезпечують стабільне навчання на великих проміжках часу, такі як, гомосинаптичні хеббівські та гетеросинаптичні гомеостатичні механізми [177], а також механізми, що змінюють структуру мережі через створення і видалення зв'язків [41]. Окремі механізми добре вивчені і побудовані математичні моделі, проте залишається проблема їх об'єднання в єдину робочу модель.

В цьому підрозділі представлена модель, що об'єднує навчання з підкріпленням та навчання без учителя з використанням гомеостатичних механізмів. Модель навчання включає дві змінні, які відслідковують силу зв'язку між нейронами на двох різних часових масштабах. Завдяки взаємодії змінних між собою, вдалося включити швидке підсилення зв'язку при надходженні сигналу підкріплення, а також формування стійких зв'язків до статистично частіших вхідних стимулів. Гомеостатичний механізм слугує формою швидкого негативного зворотного зв'язку, який протидіє насиченню сили зв'язків до граничних

значень.

1.2.1 Модель

Представлена модель подібна до існуючих підходів з використанням допоміжних слідів (eligibility traces), де додаткова змінна відслідковує нещодавню активність, і у разі надходження сигналу підкріплення підсилює зв'язок [47]. Проте, в даному випадку, змінна, що відслідковує нещодавню активність виступає сама як сила зв'язку, і тому має місце форма короткострокової пам'яті. Також, в поєднанні з повільною змінною, можливо впровадити навчання без учителя, де нейрон вибирає статистичні закономірності з вхідного сигналу.

В моделі було симульовано один вихідний нейрон, який отримує сигнал від 100 вхідних нейронів. Ваги тільки додатні, тому вхідні нейрони можуть тільки збуджувати вихідний. В якості моделі нейрона використано стандартну дисипативну модель інтегруй-активуй, що описує зміну напруги на мембрані.

$$C \frac{du}{dt} = g_{\text{leak}}(u_{\text{rest}} - u) + g_{\text{AMPA}}(t)(u_{\text{AMPA}} - u) \quad (1.7)$$

Якщо $u > u_{\text{thr}}$, то напруга скидається $u = u_{\text{rest}}$ і нейрон генерує спайк. $g_{\text{leak}}, u_{\text{rest}}, u_{\text{AMPA}}$ - константи. Провідність АМПА рецепторів g_{AMPA} описується наступним чином:

$$\tau_{\text{AMPA}} \frac{dg_{\text{AMPA}}}{dt} = -g_{\text{AMPA}} + \sum_{i \in \{\text{input}\}} w_i S_i(t) \quad (1.8)$$

Таким чином g_{AMPA} збільшується на w_i в момент надходження спайку від i -го вхідного нейрона. $S_i(t) = \sum_f \delta(t - t_i^f)$, де $\delta()$ - функція Дірака (дельта функція) і t_i^f - час надходження спайку. Зміна сили зв'язку описується наступним чином:

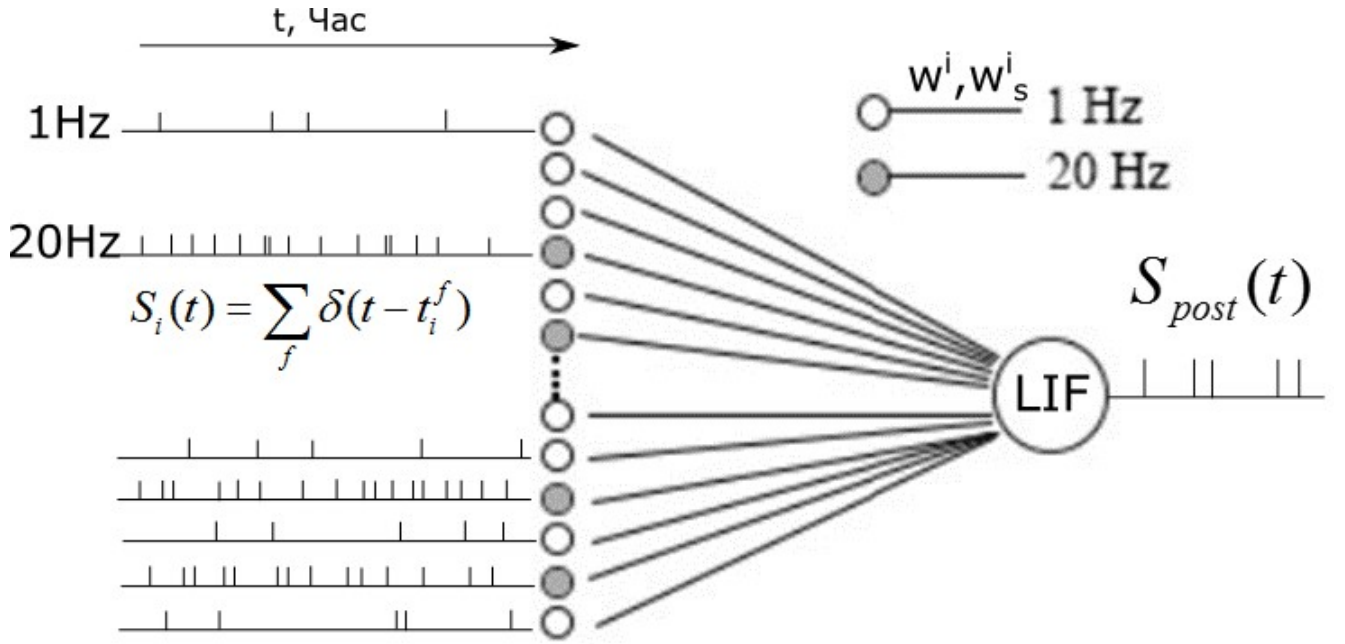


Рис. 1.8: Схематичне зображення мережі. 100 вхідних нейронів з активацією за розподілом Пуассона. 1 вихідний описується через стандартну дисипативна модель спайкових нейронів інтегруй-активуй. Подається вхідний патерн – 10-15 більш активних нейронів, позначено сірими кружечками. Кожний зв'язок описується двома змінними.

$$\begin{aligned} \frac{dw^i}{dt} &= \frac{1}{\tau_w}(w_s^i - w^i) + \frac{1}{\tau_{hebb}}(S^i(t)(x^i \cdot p_{slow}) - S_{post}(t)p), \\ \frac{dw_s^i}{dt} &= \frac{1}{\tau_{w_s}}(w^i - w_s^i)M + \frac{1}{\tau_{hom}}w_s^i(1 - \frac{y^s}{A})^3. \end{aligned} \quad (1.9)$$

Це власне і є запропонована модель, детальне пояснення якої буде дано нижче. Тут x, p_{slow} сліди спайків для вхідного нейрона і p – для вихідного. Слід розуміється як низькочастотний фільтр з параметром τ і описується наступними рівняннями:

$$\begin{aligned} \tau_x \frac{dx^i}{dt} &= -x^i + S^i(t) \\ \tau_p \frac{dp}{dt} &= -p + S_{post}(t) \\ \tau_{p_{slow}} \frac{dp_{slow}}{dt} &= -p_{slow} + S_{post}(t) \end{aligned} \quad (1.10)$$

Симуляція виконувалась наступним чином. 100 вхідних нейронів генерували випадково спайки з розподілом Пуассона з частотою $\nu^{bg} = 1Hz$. Вхідний патерн

це група з 10-15 нейронів зі збільшеною частотою активації $\nu_i(t) = \nu_i^{bg} + \mu_i(t)$, де $\mu_i(t) = 19Hz$. Схематичне зображення мережі дано на Рис 1.8.

Симуляція була виконана в мові програмування Python з використанням бібліотеки Brian2. Код програми з деталями реалізації та значеннями параметрів доступний на <https://github.com/hronoses/ProjectOne>.

Огляд моделі

Сила зв'язку між двома нейронами пов'язана з кількістю і ефективністю АМПА рецепторів, що регулюється через взаємну активність вхідного і вихідного нейронів. АМПА рецептори розміщуються на шипику, малому дендритному відростку, який приймає, обробляє і зберігає сигнал від вхідного нейрону. Всі основні збуджуючі нейрони головного мозку на місцях контакту з іншими нейронами утворюють шипики і вважається, що це є субстратом пам'яті. Визначення сили зв'язку залежить не тільки від моменту навчання, але й від динаміки стану шипика, яка має повільний масштаб часу і залежить від інших рецепторів і регуляторних білків. Тому, пропонується розрізняти між моментальним значенням сили зв'язку і динамічним станом шипика, який впливає на значення сили зв'язку і після синхронної активації вхідного і вихідного нейронів.

Також, необхідність опису зв'язку між нейронами кількома змінними виникає при навчанні з підкріпленням. Вважається, що нейрон повинен відслідковувати нещодавню активність щонайменше протягом 1-5 секунд, щоб при надходженні сигналу підкріплення лише потрібні синапси підсилили зв'язок. Для цього вводять додаткову змінну, яку в літературі називають допоміжним слідом (eligibility trace). Нещодавно встановлено, що це теоретичне припущення має місце в реальному нейроні [66]. Детальні молекулярні механізми ще мають бути виявлені, але потенційні кандидати – певні кінази і фосфатази, що регулюють силу зв'язку через модуляцію кількості та ефективності АМПА рецепторів, мають часовий масштаб залишкової активності порівняний з тим, що має до-

поміжний слід в теоретичних моделях. Тому, є підстави надати допоміжному сліду фізичної інтерпретації, що й робиться в представленій моделі.

Інша важлива риса при формуванні сили зв'язків – це стабільність. Хеббівське правило, яке каже, що зв'язки між нейронами А і Б підсилюються, якщо нейрон А бере участь в активації Б, за своєю природою має позитивний зворотній зв'язок, тобто збільшення сили зв'язку призводить до збільшення активації, яка в свою чергу ще більше збільшує зв'язок. В біологічних мережах стабільність забезпечується завдяки багатьом механізмам, через збалансовану активацію пригнічуючих нейронів [3], коротко-часове пригнічення чи швидку форму гомеостазу [27, 58, 75]. В математичних моделях хеббівське правило має бути доповнене стабілізуючими механізмами, і як було нещодавно показано, таке доповнення, веде до формування асоціативної пам'яті в спайкових нейронних мережах [177].

Підсумовуючи, запропонована модель базується на кількох ідеях:

- Є дві змінні, що описують силу зв'язку між двома нейронами: швидка і повільна. Ці змінні взаємодіють між собою на різних часових масштабах.
- Повільна змінна описує стан шипика і може швидко зрівнятися зі швидкою, при надходженні сигналу підкріплення.
- Нейрон повинен мати частоту активації в певних межах. Гомеостатична пластичність забезпечує цю стабільність через масштабування сил зв'язку.

Модель описується наступними рівняннями:

$$\begin{aligned}\frac{dw}{dt} &= \frac{1}{\tau_w}(w_s - w) + \frac{1}{\tau_{hebb}}Hebb \\ \frac{dw_s}{dt} &= \frac{1}{\tau_{w_s}}(w - w_s)M + \frac{1}{\tau_{homeo}}Homeo\end{aligned}\tag{1.11}$$

w - швидка змінна, миттєве значення сили зв'язку, w_s - повільна змінна, описує стан шипика. M – сигнал підкріплення, $M = 1$ без сигналу. $Hebb$ та $Homeo$ описують гомосинаптичну хеббівську пластичність та гетеросинаптичну гомеостатичну пластичність відповідно. Також повинно мати місце співвідношення

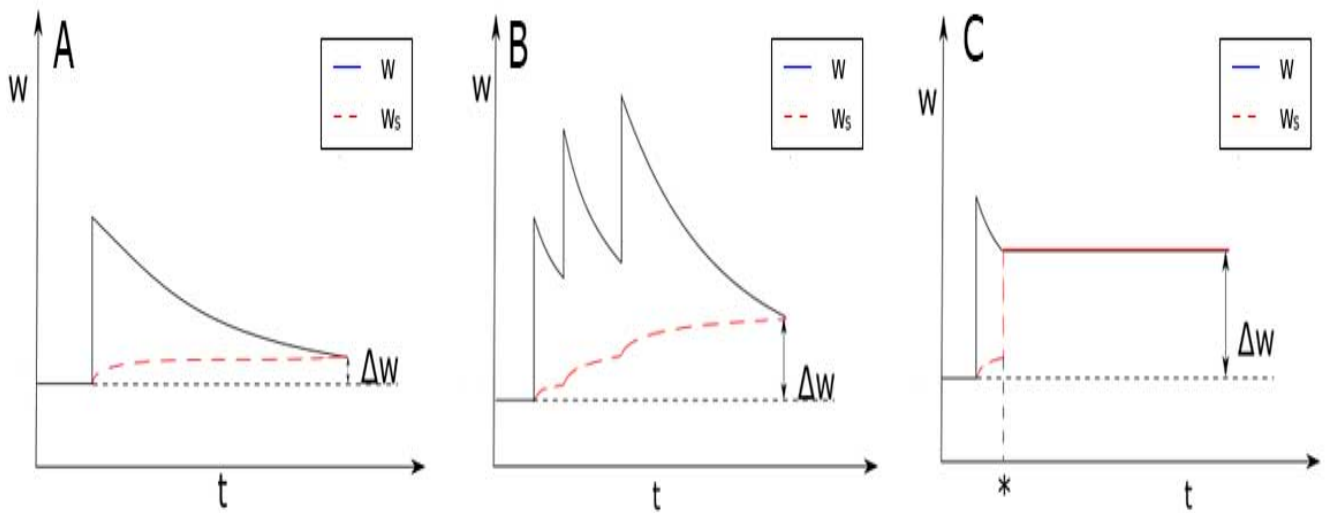


Рис. 1.9: Зображення еволюції швидкої та повільної змінних. А: зображення динаміки двох змінних при одному моменті навчання. В: повторення навчання, призводить до збільшення сили зв'язку. С: сигнал підкріплення, майже миттєво збільшує силу зв'язку

часових масштабів $\tau_{w_s} \gg \tau_w$, що визначає відношення швидкості динаміки двох змінних.

Розглянемо один момент навчання, коли відповідно до хеббівської пластичності змінна w зростає (Рис 1.9А). Тоді, згідно рівнянь 1.9, w починає зменшуватись, а w_s зростає. Загалом, повільна змінна збільшиться на Δw , що залежить від співвідношення $\frac{\tau_w}{\tau_{w_s}}$. Якщо навчальна подія повториться, тоді загальний ріст буде більшим (Рис.1.9В). Це добре відповідає твердженню, що реактивація навчання покращує пам'ять [122].

Основна відмінність запропонованої моделі навчання в тому, що вона природно включає навчання з підкріпленням. Припустимо, що після навчання нейрон отримує сигнал підкріплення M . Тоді швидкість зміни ваги w_s буде значно більша вона наближається до $w_s \rightarrow w$, що виразиться у великій зміні швидкої ваги Δw (Рис.1.9С).

Навіть без явного виразу для хеббівської та гомеостатичної пластичності, ця модель показує цікаві властивості:

- Сила зв'язку залежить не тільки від моменту навчання, але й від довшого часового проміжку.

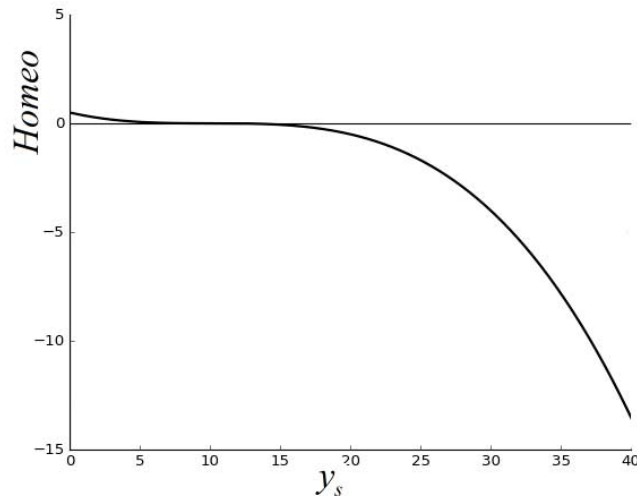


Рис. 1.10: Графік зміни сили зв'язку в залежності від сліду активації вихідного нейрона

- Модель одночасно включає навчання без учителя і навчання з підкріпленням

Для хеббівської форми використано правило триплетної спайк-часово-залежної пластичності [133], яка може бути зведена до поширеного правила ВСМ [54].

Для гомеостатичної пластичності використано наступний вираз:

$$Homeo = w_s \left(1 - \frac{y_s}{A}\right)^3 \quad (1.12)$$

w_s — відповідає мультиплікативній природі пластичності, коли зміна сили зв'язку залежить від самої величини. y_s — слід активації нейрона $\tau_y \frac{dy_s}{dt} = -y_s + S_{post}(t)$, A — задана частота активації.

Кубічна залежність (рис.1.10) була використана для того, щоб зробити від'ємний зворотний зв'язок сильнішим при сильній активації і слабшим коли частота активації близька до заданої частоти. Але, недоліком такого виразу є те, що зміна сили зв'язку при низькій активації значно повільніша і у випадку $w_s = 0$ зміна відсутня зовсім.

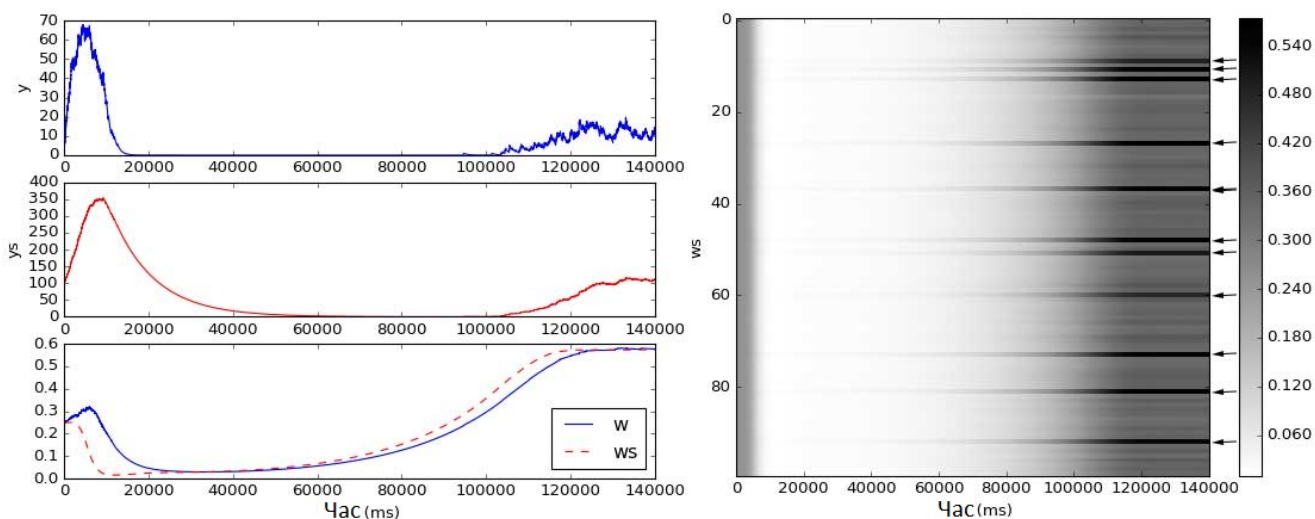


Рис. 1.11: Регуляція активності нейрона. Зліва: Верхні два графіка відповідають сліду активності активації постсинаптичного нейрона з часовими сталими 1с та 10с. Знизу графік показує залежність від часу двох змінних зв'язку до одного нейрона з групи, які активувались від патерну. Справа: Залежність повільної змінної з часом для всіх 100 зв'язків. Стрілками вказано активний патерн. Градієнт сірого вказує на роботу гомеостатичної пластичності

1.2.2 Аналіз та результати

Далі наводяться результати трьох експериментів із спайковою мережею та запропонованою моделлю пластичності, з якими також можна ознайомитися в [128].

Дослідження стабільності зв'язків

100 вхідних нейронів надсилають спайки з розподілом Пуассона з частотою 1Hz до одного вихідного нейрона. На мережу перші 20 секунд був поданий активний патерн, в результаті чого 11 вхідних нейронів активувалися з частотою 20 Hz. Відповідно до правила навчання 1.9, зв'язки пов'язані з активним патерном підсилювалися.

Зі збільшенням частоти активації вихідного нейрона, член *Homeo* зменшує зв'язки і нейрон перестає збуджуватись. Тоді гомеостатична пластичність починає працювати в іншу сторону для вирівнювання активації до бажаного рівня. Завдяки тому, що в 1.9 входить w_s , навчений патерн збільшує зв'язки сильні-

ше, так як в нього більше значення w_s і тому можливе виділення і збереження навченого патерну. Така властивість називається мультиплікативністю і узгоджується з експериментальними даними на реальних нейронах [58]. У випадку відсутності мультиплікативності, навчений патерн після відновлення активації нейронна був би відсутній. Отримані результати, коли при навчанні сили зв'язку збільшується, а потім від гомеостатичної пластичності зменшуються, схожі на отримані в [175], проте в цьому випадку кінцеве значення сили зв'язку навченого патерну більша за початкове.

Також, варто відмітити, що стабілізаційний ефект виникає ще з різниці $w_s - w$ у рівнянні 1.9. Чим більша різниця, тобто чим далі від стабільності, там більше зменшується вплив хеббівського члена, який відповідає за виникнення нестабільності.

Якщо врахувати, що спонтанна активність, яка симулюється через розподіл Пуассона, насправді має просторово часові кореляції [101], які відповідають попередньо навченим стимулам, тоді відновлення гомеостазу відбудеться швидше, і не буде такого довгого періоду неактивності.

Тут застосовано тільки один з багатьох механізмів стабільності. Для повної картини, нейрони інгібітори повинні бути додані. Проте, навіть представлене просте правило гомеостатичної пластичності узгоджує значення ваг відповідно до вхідної активності і протидіє насиченню сили зв'язків до граничних значень $(0, w_{max})$, і головне – зберігає вивчений патерн.

Вплив активації кількох патернів

Далі досліджено поведінку моделі у випадку активації кількох патернів. Головне питання було чи збережеться сила зв'язків до кількох патернів при гомеостатичній регуляції і чи буде виділятися частіший стимул. Було подано на вхід 5 різних патернів, де кожен активує 10-15 нейронів; один патерн був частіший інших.

Як видно з Рис 1.12 частіший патерн отримав найбільший приріст сил

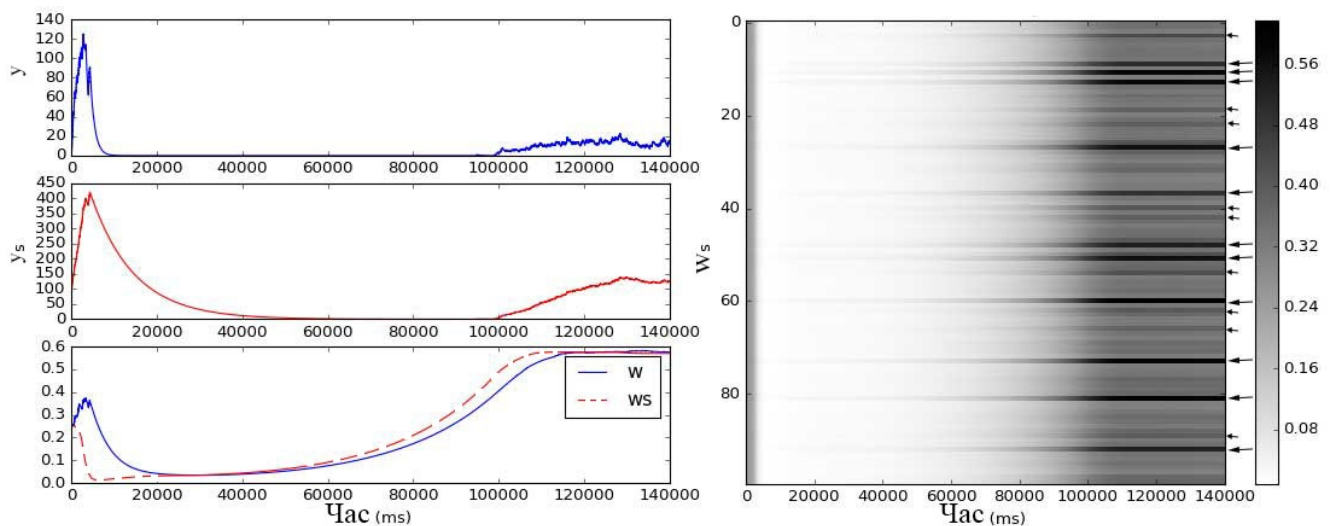


Рис. 1.12: Поведінка моделі при активації кількох патернів. Великі стрілки вказують на частіший патерн. Малі стрілки вказують на інші менш частіші патерни, які зберегли своє представлення.

зв'язку, як і очікувано. Також, збережено інші патерни, хоча і з меншим значенням сили зв'язку. Таким чином, нейрон навчається на більш статистично вагомій стимулі, тим самим реалізуючи простий варіант навчання без учителя.

Варто зазначити, що процедура масштабування ваг згідно виразу 1.12 схожа до нормалізації ваг $\|w\| = (\sum_i w_i^p)^{1/p} = 1$, що використовується в правилі Оджі [124], і тому дана модель ймовірно також виділяє принципові компоненти з вхідного стимулу, але для підтвердження потрібні додаткові дослідження.

Дослідження навчання з підкріпленням

В наступному обчислювальному експерименті також подавалися на мережу кілька патернів, але в цьому випадку вони комбінувалися з сигналом підкріплення.

Два патерни були подані на початку протягом 5-ти секунд роботи мережі. Далі, в 9.8с був поданий третій патерн і 0.3с після, наданий сигнал підкріплення. Відповідно до моделі, повільна змінна зросла до значення швидкої, в результаті чого відбулося швидке навчання. Якщо у випадку навчання без учителя потрібно надавати велику кількість раз потрібний патерн, щоб нейрон вважав його

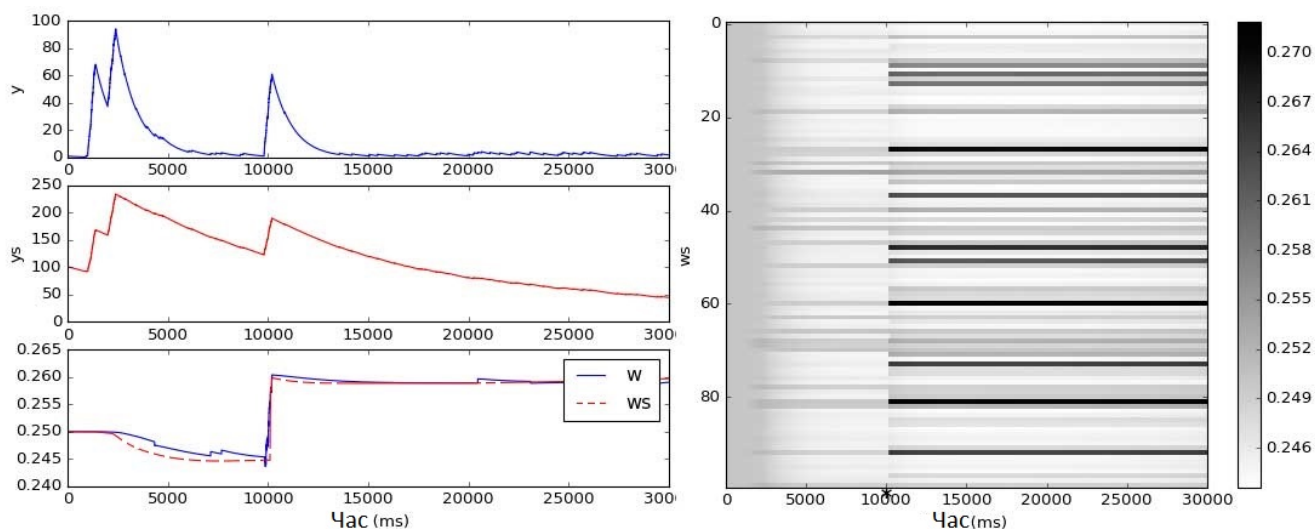


Рис. 1.13: Дослідження підкріплення ваг зв'язків. Зліва: Загалом було подано три патерни і останній підкріплено. Справа: Зірочка вказує на час надходження сигналу підкріплення. Підкріплений патерн має значно більші значення синаптичних ваг, порівняно з іншими, попередніми патернами

важливим і навчився до певного рівня, то у випадку навчання з підкріпленням всього одного разу достатньо, щоб відбулося значне перерозподілення ваг.

1.2.3 Обговорення

Основна відмінність запропонованої моделі у використанні двох змінних для опису сили зв'язку. Через взаємодію змінних вдалось зручно впровадити навчання з підкріпленням зі збереженням навчання без учителя. Через те, що змінні мають різні часові масштаби, швидка змінна реагує на вхідний стимул, а повільна з часом консолідується, що в свою чергу допомагає в стабілізації зв'язку.

При симуляції моделі на простій нейронній мережі, показано здатність до вивчення активного патерну, збереження сили зв'язку у випадку багатьох вхідних патернів, а також здатність до стабілізації зв'язків для підтримання заданого рівня активації нейрона. Також, показано здатність моделі до швидкого навчання через сигнал підкріплення.

В роботі було обрано стала часу швидкої змінної w порядку секунди, і тому вона відслідковувала нещодавню активність і використовувалась для навчання

з підкріпленням. Проте, якщо взяти часову константу порядку десятків хвилин, тоді швидку змінну можна інтерпретувати як опис ранньої фази довготермінової пластичності, а повільну як консолідацію сили зв'язку. В цьому випадку сигналом підкріплення може виступати висока активність вихідного нейрона, який генерує пачки спайків, і тоді швидка консолідація співвідноситься з експериментами у рамках гіпотези синаптичної мітки і закріплення [146].

Представлена модель схожа на загальний підхід до навчання з підкріпленням на спайкових нейронних мережах – допоміжні сліди [47]. В таких моделях вводиться допоміжна змінна, яка відслідковує активність і при необхідності підкріплює зв'язок. Проте в запропонованому підході швидка змінна грає роль сили зв'язку і може слугувати формою короткострокової пам'яті, яка впливає на активацію нейрона.

Модель має цікаву можливу біологічну інтерпретацію. Після активації синапсу в постсинаптичний нейрон входять іони Ca^{2+} , які активують регуляторні протеїни, такі як CaMKII та PKA, які в свою чергу спричиняють швидке переміщення АМПА рецепторів зі внутрішніх сховищ на поверхню мембрани [159, 39]. Активація протеїнів спадає протягом хвилини [99] і нестійкі АПМА рецептори інтерналізуються назад в клітину. Якщо прийняти гіпотезу про те, що АМПА рецептори формують стабільний кластер на поверхні мембрани [153], тоді зміну повільної змінної можна інтерпретувати як включення чи виключення рецепторів в кластер. У випадку надходження сигналу підкріплення активуються допамінові рецептори, які запускають невідомі молекулярні механізми, які закріплюють вільні АМПА рецептори в кластери, тим самим збільшуючи силу зв'язку. Тому, модель може мати біологічну основу, де w_s відповідає розміру шипика і молекулярній машинерії, що підтримує його стан, і w відповідає кількості та ефективності АМПА рецепторів.

В основі роботи нейрона лежать багато механізмів пластичності, що забезпечують навчання просторово-часових стимулів і в результаті кооперації з іншими нейронами формується пам'ять складних вхідних стимулів. Об'єднання всіх не-

обхідних механізмів в єдину модель лише на початку шляху, але вже зараз є цікаві результати [177]. Навчання без учителя і з підкріпленням разом з механізмами стабільності реалізовані у спайковій нейронній мережі має перспективи перерости в ефективну обробку інформації і прокласти шлях до адаптивних роботів.

1.3 Проблеми спайкових нейронних мереж

Наукові роботи частіше всього пишуть безособово, так як опис природи не має залежати від людини, і наступні покоління використовуватимуть ці знання в чистому вигляді для вирішення нових задач. Проте, особисті історії допомагають зрозуміти походження ідеї та мотивації, чому саме ця задача важлива.

Після першого року аспірантури, коли я вже досить добре вивчив обчислювальну нейронауку, і розробив модель навчання, представлену вище, я проходив стажування у французькій лабораторії спеціалізованій на робототехніці. Захотів перевірити на практиці методи обчислювальної нейронауки, чи може спайкова нейронна мережа контролювати робота та навчатися від сенсорного входу. Мене чекало розчарування. По перше, правила навчання, такі як STDP, не застосовні на практиці. Хоча й останні роботи показують формування стабільних слідів пам'яті, та все одно стоїть проблема навчання послідовностей, зокрема, навчання тривалості та інтервалів. Були статті, на намагалися будувати мозкоподібних роботів (brain-based devices) на основі спайкових нейронних мережах [94, 37, 44], але ці роботи мали обмежені можливості розпізнавання, де часто багато чого задавалося вручну. По друге, стоїть обчислювальне обмеження. На потужних персональних комп'ютерах можна симулювати в режимі реального часу спайкову нейронну мережу розміром не більше ніж 10 000 нейронів. Якщо використовувати більш складніші моделі нейрона чи правила навчання кількість нейронів ще менша. І тому важко побудувати систему контролю робота з навчанням. Щоправда, є можливість використовувати суперкомп'ютери

в рамках Human brain project Robotics platform [64], проте залишається перша проблема. Також, перспективним напрямком вважається нейроморфні мікросхеми, наприклад, IBM TrueNorth [7], які значно пришвидшують симуляцію велико-масштабних нейронних мереж, але перша проблема нікуди не зникає.

В обчислювальній нейронауці досі стоїть проблема вибору необхідного масштабу опису. Чи потрібно враховувати допорогові флуктуації напруги, чи електричні властивості дендритів, чи, можливо, потрібно робити опис на молекулярному рівні, виділяючи концентрації різних видів молекул, кожна зі своєю роллю, чи навіть варто все моделювати на генетичному рівні? Це типова проблема висхідного підходу для складних систем, якого рівня опису буде достатньо, щоб відбувалася самоорганізація і виникали нові властивості для системи вцілому.

Взяти, наприклад, високо цитовано роботу 2015 року про формування та зчитування пам'яті у спайкових нейронних мережах [177]. В ній використовується феноменологічна модель спайкового нейрона з кількома формами пластичності: триплетна спайк-часово залежна пластичність, гомеостатична пластичність, трансмітер-викликана пластичність. Модель лише одного нейрона описується більше 10-ма диференціальними рівняннями, з більше ніж 20-ма параметрами. Нейронна мережа розміром 4000 нейронів записує 4 патерни, тобто створює автоасоціативну пам'ять, коли частковий показ патерну відтворює весь записаний. Проте, цю саму задачу вирішує, давно створена, мережа Хопфілда набагато ефективніше з більшою ємністю, меншою кількістю параметрів та часом симуляції. Ставиться питання, навіщо така надлишкова складність, якщо не сподіваєшся на появу нових ефектів, а вирішуєш класичну задачу асоціативної пам'яті?

Тим не менш, невдача із застосування спайкових нейронних мереж відкрила інші напрямки, я вважаю більш перспективнішими. Такі області як моделі асоціативної пам'яті, теорія інформації та стиснення даних працюють з абстрактними принципами обробки інформації, часто в розподілених структурах, ти-

пу нейронних мереж. Такий "чистий" опис задачі розкриває суть обчислень, без надлишкових біологічних деталей, але й може бути біологічно подібним, тобто, коли розв'язок задачі має можливу біологічну імплементацію.

На сьогодні область обчислювальної нейронауки рухається наосліп, експериментуючи з пластичністю, яка б відповідала біологічним деталям. Тому, далі в цій роботі розглядаються більш абстрактні рішення задач формування пам'яті. Відкидаються багато біологічних деталей і формулюються спрощені математичні моделі.

1.4 Висновки

В даному розділі представлений короткий, але широкий огляд нейронауки. Показані останні дані, гіпотези та теорії організації та функціонування біологічних нейронних мереж. Також, описано обчислювальні підходи, які намагаються формалізувати біологічні дані в обчислювальні моделі. Загалом можна виділити:

1. На основі огляду нейронауки, що показує надзвичайну складність реальної нейронної мережі, аргументовано необхідність пошуку абстрактних принципів обробки інформації в розподілених структурах. Проте, біологічні дані залишаються дуже важливими, так як спрямовують та допомагають знайти рішення багатьох задач розпізнавання та кодування.

2. Взаємозв'язок повільної та швидкої змінної в запропонованій двомірній моделі пластичності дозволяє навчати стабільні зв'язки, виділяти статистично значимі патерни, та включає сигнал підкріплення. Модель, можливо, відповідає взаємодії молекулярних механізмів в шипику для стабілізації АМПА рецепторів.

Результати експериментальних досліджень даного розділу наведено в публікації [128]:

Осауленко В.М. Двовимірна модель навчання у спайкових нейронних мере-

жах з гомеостазом та навчанням з підкріпленням, System Research Information Technologies, 2017, №2, 130-140, DOI: 10.20535/SRIT.2308-8893.2017.1.12 (Фахове видання)

Розділ 2

Просторова асоціативна пам'ять

Адаптація – одна з найважливіших рис біологічних нейронних мереж. Вона проявляється через зміну сил зв'язків та зміну самих зв'язів (видалення старих та додавання нових) під впливом інформації із зовнішнього світу, що сприймається органами чуття. Її результатом є представлення довільної інформації через просторово-часові патерни активації нейронів. Наприклад, жовтий колір, напівмісячна продовгувата форма та запах мають свої представлення у відповідних частинах мозку, і все разом формує цілісне представлення об'єкту, що називається банан. Сама назва "банан і її звучання, і всі спогади представлені окремими популяціями активних нейронів.

Просторова асоціативна пам'ять це представлення тих образів, які не змінюються в часі, наприклад, колір чи форма. Ця пам'ять ґрунтується на просторових регулярностях в природі, як наприклад, лінії (ймовірність, що біля двох чорних пікселів буде третій розташований по прямій, значно більша, ніж будь-яке інше розташування). На вищому рівні, ймовірність появи жовтого кольору з продовгуватою формою, значно більша ніж з довільною формою. Тому, цей розділ присвячений тому, як формуються зв'язки, які зв'язують регулярності та формують асоціативну пам'ять.

Варто одразу зазначити, що окрім просторових, значну роль відіграють часові регулярності. Типовим прикладом є прямолінійний рух: яблуко, яке відпустити, впаде на землю по прямій лінії. В загальному, часові регулярності спричинені фізичними причинно-наслідковими законами, і зберігаються просторово-часовою пам'яттю, яка детальніше описана в третьому розділі.

Перші моделі просторової асоціативної пам'яті досліджувалися ще з 1960-х років. Проте, нещодавні дослідження біологічних нейронів, зокрема, дендри-

тних обчислень, пропонують інший підхід до зв'язування інформації. Нелінійна сумація вхідних імпульсів виконує роль детектора збігів і нейрон активується тільки тоді, коли зустрічається правильна комбінація вхідних активних нейронів. Така селективність має ряд обчислювальних переваг, що і покладено в модель представлений в даному розділі.

2.1 Постановка задача та основні підходи

Людська пам'ять побудована на принципі створення асоціацій [50], коли два і більше об'єктів або подій з'єднуються разом, таким чином, що презентація одного елементу викликає з пам'яті інші. Така пам'ять використовується живими організмами і принципово відрізняється від адресної в сучасних комп'ютерах. Адресна зберігає інформацію у вигляді "ключ:значення", де ключем виступає адреса комірки в якій збережені дані. В асоціативній – ключем виступає інше значення.

Класичний метод записати будь-яку інформацію, чи то слова, зображення, звук, чи числові дані, починається з вибору алфавіту, базових елементів з яких складається інформація. Нехай, алфавіт складається з N елементів, тоді створюється словник, де кожному елементу присвоюється номер в двійковій системі. Розмір словника дорівнює, щонайменше, $\log_2(N)$, саме стільки потрібно 0-ів або 1-ць для присвоєння унікального номеру кожному елементу. Записується інформація в бінарному вигляді в пронумеровані комірки певного розміру.

В живих організмів немає словника, їх принципи кодування інформації ще не до кінця зрозумілі. Можливо, саме тому асоціативна пам'ять не знайшла такого широкого застосування. Далі наведені основні відмінності від адресної пам'яті:

- ключем виступає інше значення. Наприклад, для зчитування коду жовтого кольору потрібно представити код для форми банана і навпаки;
- асоціативна пам'ять добре працює при шумі. Якщо вхідне значення ви-

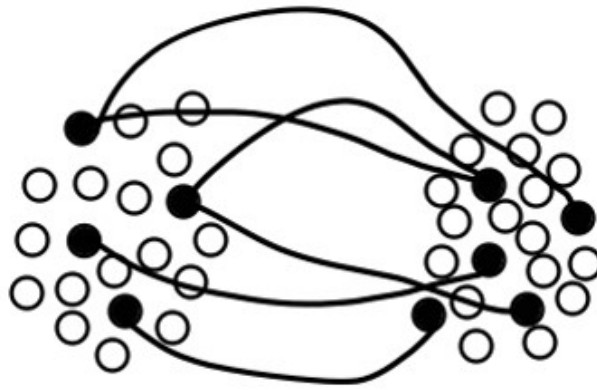


Рис. 2.1: Ідея навчання як створення асоціативних зв'язків між активними нейронами

кривлене, все одно, буде видобуто правильну асоціацію. В адресній пам'яті за невірної адреси, потрібно перебрати послідовно всі існуючі і знайти найбільш схожу до викривленої;

- асоціативна пам'ять природно впроваджується на паралельні архітектури, де працює значно швидше адресної;

- обчислення не відділені від пам'яті, тобто, щоб видобути спогад, потрібно активувати нейрони, що їх кодують.

Важливим, і досі до кінця не вирішеним питанням є представлення даних. В якій формі представляти звуки, слова, зображення. Дана робота слідує підходу представлення даних в розріджених бінарних векторах. Тому, створення асоціації між жовтим кольором і бананом, означає створення асоціації між двома векторами. З точки зору біології, зв'язати два вектори означає створити асоціативні зв'язки між активними нейронами. Схематично, це зображено на рис. 2.1.

Математично, задача асоціативної пам'яті ставиться як пошук відображення $f : \mathbf{x} \rightarrow \mathbf{y}$ що з'єднує пари векторів $\{\mathbf{x}^\mu \rightarrow \mathbf{y}^\mu, \mu = 1 : R\}$. R – кількість записаних пар, $\mathbf{x}^\mu \in \{0, 1\}^N$, $\mathbf{y}^\mu \in \{0, 1\}^M$ бінарні вектори розмірності N та M відповідно. При цьому, презентація вхідного, або його ще називають адресного, вектору $\tilde{\mathbf{x}}^\mu$ повертає асоційований (з \mathbf{x}^μ) вектор \mathbf{y}^μ , або найбільш близький до записаного $\tilde{\mathbf{y}}^\mu$. Вводиться критерій допустимості ϵ , який визначає допустиме

відхилення видобутого \tilde{y}^μ від записаного y^μ . Значення R_{max} визначається як максимальна кількість збережених пар при якій помилка рівна ϵ .

Моделі Вілшоу, Хопфілда та варіації

Якщо вхідний і вихідний вектор різні $x \neq y$, то пам'ять називається гетеро-асоціативною, і класична модель в цьому випадку – модель Вілшоу. Якщо вектори однакові $x = y$, то пам'ять авто-асоціативна, і класична модель - модель Хопфілда. Моделі двох типів знаходяться під впливом моделі нейрона Маккалоха і Піттса та правила навчання Хебба. Тому, зазвичай використовується зважена сумація вхідного вектору $y_j = f(\sum_i w_{ij}x_i)$ та навчання як збереження парних кореляцій $w_{ij} = g(x_i \cdot y_j)$.

Модель Вілшоу - перша робота (відома автору дисертації), яка запропонувала обчислювальну модель асоціативної пам'яті [172]. Представлення даних у вигляді бінарних векторів зі значеннями $\{0, 1\}$. Функція активації – зважена сумація вхідного вектору та порівнянням з порогом: $y_i = \{1, \text{якщо } \sum_j w_{ij}x_j \geq \theta; 0, \text{інакше}\}$. Функція навчання – хеббівське правило з бінаризацією: $w_{ij} = \min(1, \sum_{\mu=1}^R x_i^\mu y_j^\mu)$. Ті нейрони, які активні разом $x_i^\mu = 1, y_j^\mu = 1$ створюють бінарні зв'язки між собою $w_{ij} = 1$. Важливим є вибір порогу θ , так як від нього залежить якість видобування. В оригінальній роботі було запропоновано використовувати поріг величиною в кількість активних нейронів $\theta = a$ (тут і далі, активні нейрони, те ж саме, що й елементи вектора зі значенням "1"). Пізніше, в роботі [56] було показано вищу ефективність використання процедури kWTA (k-winners-take-all), при якій k нейронів з найбільшою активацією (зваженою сумою) стають активними $y_i = 1$, а всі інші зануляються. Це еквівалентно використанню порогу зі значенням активації k -того відсортованого нейрона. Також, процедура kWTA покращує результати з частковою зв'язністю, коли не всі зв'язки w_{ij} можливі, та зі спотвореними вхідними векторами \tilde{x} . Було показано, що ємність пам'яті максимальна за розріджених векторів, в яких кількість

активних нейронів (a одиничних елементів вектору) набагато менша за всю популяцію (розміру вектора N , тобто $a \ll N$). Розрідженість визначається як відношення $s = a/N$.

Інший підхід, автоасоціативної пам'яті, реалізовує модель Хопфілда [70]. Вхідний та вихідний вектор однакові і стоїть задача відтворення зашумленої версії (noise correction). Вхідні вектори не обов'язково є розрідженими, і значення активації вже не $\{0, 1\}$ як в моделі Вілшоу, а $\{-1, 1\}$. Використовується та ж активаційна функція з порогом $\theta = 0$, так як значення активації нейронів можуть бути від'ємні. Також реалізується подібний хеббівський механізм навчання $w_{ij} = \sum_{\mu=1}^R x_i^{\mu} y_j^{\mu}$, тільки ваги вже не бінарні. Такі правила навчання дозволяють успішно записувати в мережу патерни активації (тобто вхідний вектор з певним розподілом одиниць). Відтворення відбувається за кілька кроків з поступовим наближенням до атрактора (вихідного патерну, який може бути рівний вхідному або помилковому). При записі занадто багатьох патернів можуть бути помилкові відтворення і тоді на виході видобувається помилковий (spurious) патерн: або інший збережений або зовсім новий. Теоретичний аналіз моделі ставить на меті визначити за яких параметрів мережі (a, N, w) можна записати максимум патернів з прийнятним рівнем помилок.

Обидві моделі вирішують одну і ту ж проблему – пошук найближчого сусіда (best match problem, з поміж N елементів знайти елемент з найменшою відстанню від заданого). Однак їх алгоритми побудовані на нейронних мережах, тому вони природно можуть бути обчислені в паралельний спосіб. В цьому і полягає важлива відмінність від звичайної адресної пам'яті.

Максимальна кількість записаних асоціацій R_{max} визначається за допомогою критерію надійності ϵ , який показує допустимий рівень помилок. Зазвичай, обирається $\epsilon = 0,01$ (проте залежить від області застосування), що означає, що пам'ять може припуститися не більше 1% відсотку помилок. Тобто, якщо зберігається 10000 патернів, і при менше ніж 100 відтворених невірних патернів

модель вважається робочою.

В моделі Хопфілда при випадкових патернах активації максимальна кількість збережених патернів в асимптотичному випадку прямує до $0,14N$ при $N \rightarrow \infty$, де N розмір мережі [8]. Для моделі Вілшоу ця характеристика прямує до $0,69N$ [89]. Проте, так як в моделі Хопфілда використовується щільна активація, тобто в середньому половина нейронів активні, тому й в ній передається більше інформації. Ємність каналу враховує цю особливість, тому мережева ємність мережі Хопфілда дорівнює 0.72 порівняно з 0.69 в мережі Вілшоу.

Ці дві класичні моделі схожі, та модель Хопфілда викликала значно більший інтерес у спільноти, хоч і з'явилася на 13 років пізніше. Пізніше виникло багато варіацій моделей асоціативної пам'яті, які намагаються покращити три основні критерії наведені вище. Двонаправлена (bidirectional) асоціативна пам'ять створює зворотні зв'язки $\mathbf{y} \mapsto \mathbf{x}$ та ітеративно обраховує активацію не тільки для $\mathbf{y}(t)$, але й для вхідного вектору $\mathbf{x}(t)$, яка залежить від часового кроку [92]. Така активація сходиться до правильного вектору і нагадує біологічні нейронні мережі, що мають багато зворотніх зв'язків. Також, проводилися дослідження по асоціативній пам'яті в біологічно подібних спайкових нейронних мережах [158, 74, 4], більше в огляді [57]. Подібно до двонаправленої моделі, в таких мережах потрібний патерн видобувається з часом. Варто відзначити роботу [177], згадану в першому розділі, в якій зібрано модель, що враховує основні обчислювальні принципи в спайкових нейронних мережах, від складної функції активації нейрона, що враховує попередню історію активації, до поєднання триплетної спайк-часово-залежної пластичності з принципами гомеостазу. В цій роботі показано здатність мережі до створення асоціацій на прикладі візуальних образів.

Останнім часом особливий інтерес звернули на структурну пластичність в моделях асоціативної пам'яті [90]. Структурна пластичність відповідає не за зміну сили зв'язку, як у правилі Хебба, а за зміну самих зв'язків, тобто додавання та видалення. Структурна пластичність відіграє ключову роль в на-

вчанні тварин, створюючи нові зв'язки між корельованими популяціями. І вже показано переваги такої пластичності в штучних моделях [90]. Динамічне перев'язування зв'язків дозволяє найбільш ефективно розподіляти обмежені ресурси під конкретну задачу.

Також були запропоновані мережі з синапсами вищих порядків, в кінці 80-х та на початку 90-х років, натхненні новими даними з нейронауки. Було показано, що кілька синапсів можуть групуватися локально на дендритному сегменті і це покращує кількість збережених патернів [132, 26, 2, 9, 53, 97, 165]. В таких моделях зв'язки навчаються відслідковувати кореляції вищих порядків, наприклад, активність вже трьох нейронів $x_i x_j x_k$. Нещодавно відродився інтерес до таких нейронних мереж [96] в силу нових даних нейронауки в області нелінійної дендритної інтеграції. І хоча, такі нейронні мережі мають величезну кількість максимально записаних патернів, проте через експоненційно зростаючу кількість зв'язків, вони обчислювально надто складні. Наприклад, для навчання потрійних кореляцій в повнозв'язній мережі, потрібно N^3 кількість зв'язків.

На подібних принципах побудований сігма-пай нейрон, який використовує активаційну функцію точного детектора збігів $y = \prod_i w_i x_i$. У бінарному випадку, якщо всі вхідні нейрони активні одночасно $x_i = 1$ та їх ваги ненульові, тільки тоді вихідний нейрон активний. Раніше було показано можливість використання нейронів сігма-пай для створення асоціацій [134]. Пізніше, було показано, що врахування дендритних дерев збільшує ємність "нервової тканини"[137]. Проте, ще не було спроб (до відома автора дисертації) побудувати та теоретично дослідити асоціативну пам'ять подібно до моделі Вілшоу. Таке дослідження представлено в даній роботі.

Кодування інформації

Перш ніж перейти до моделі, варто обговорити кодування інформації або представлення даних. Розрізняють три випадки векторного представлення: локаль-

не кодування, щільне кодування та розріджене.

При локальному - в мережі з N нейронів активується лише один, який відповідає одному з N об'єктів з середовища. При такому кодуванні легко можна створювати асоціації, так як немає перекриття різних патернів. Проте, такий спосіб вкрай неефективний, так як всього можна записати N об'єктів.

Протилежний випадок - це активація багатьох нейронів, в середньому, половини всіх активних клітин. За такого кодування досягається максимальна ентропія мережі, а отже передається максимум інформації. Тим не менш, на практиці різні патерни мають занадто сильне перекриття (в середньому 50 % активних клітин спільні) і лінійна сумація та збереження парних (хеббівських) кореляцій не дозволяють розділити патерни і асоціативна пам'ять часто видає помилкові результати.

Компромісним рішенням між двома екстремальними випадками є розріджене кодування. При кодуванні з малою ентропією лише мала кількість нейронів активна в певний момент, яка і кодує певний об'єкт з середовища. За такої умови немає великого перекриття різних патернів, і тому можливе їх розділення [11]. Також, розріджене кодування забезпечує більшу ємність збережених патернів в асоціативній пам'яті [148]. Для моделі Вілшоу, при рівні активності $a = \log_2 N$ досягається максимальна ємність [172]. Також, численні експериментальні дані показують, що в нейронних мережах в областях, що відповідають за асоціативну пам'ять (гіпокамп, деякі області базальних ганглій та поверхневі шари кори головного мозку), спостерігається розріджена активація мережі. Проте, залишається ще відкритим питанням як розріджене кодування формується від активації сенсорних клітин [143, 144]. Одним із способів запропонований в роботах [45, 29] базується на навчанні через конкуренцію між нейронами, що також використовується в алгоритмах квантизації векторів [69]. Інакший варіант використовувати алгоритм випадкової проекції у вищу розмірність із застосуванням процедури kWTA, яка залишає активним k клітин. Нещодавно було показано, що в нюховій нейронній мережі мух відбувається саме така процедура [33], що

подібно до локально чутливого хешування (local sensitive hashing), та дозволяє мухам робити пошук найближчих спогадів від певного запаху.

З перевинайденням алгоритму зворотного поширення похибки в 1986 році напрямок навчання нейронних мереж як оптимізація певної функції став домінуючим. В глибоких нейронних мережах навчання end-to-end і проблема кодування даних не виникає. Це одна з причини, чому інтерес до моделей асоціативної пам'яті, які використовують більш біологічно подібні локальні правила навчання зменшився. Проте, нещодавно, почав розвиватися напрямок додавання пам'яті в класичні нейромережі для вирішення проблеми катастрофічного забування [174]. Мережа перед навчанням від нових даних зберігає свій стан в іншій нейронній мережі, тому моделі асоціативної пам'яті можуть поєднатися з напрямком deeplarning в майбутньому.

Хоча, класичні моделі асоціативної пам'яті існують вже десятки років, і добре теоретично описані, все ще залишається ряд проблем. Зокрема, проблема досягнення хорошої ємності для малого розміру мережі, а не для асимптотичного випадку. Також, проблема роботи мережі з щільним кодуванням. Залишається невирішена проблема формування розрідженого представлення від сенсорних клітин. В реальних задачах патерни для асоціювання далекі від випадкових і часто сильно корельовані, що суттєво обмежує ємність пам'яті. Також, цікавим невирішеним напрямком є кодування ймовірнісних відношень між об'єктами, та відтворення обчислень подібних до баєсівських в нейронній мережі.

В наступному підрозділі пропонується підхід заснований на нових біологічних деталях обчислень в нейроні, зокрема дендритному дереві. Показано, що створення асоціативної пам'яті на основі моделі нейрона сігма-пай, можна вирішити деякі з перелічених вище проблем.

2.2 Асоціативна пам'ять на основі дендритних обчислень

Дендритні обчислення

Як уже було сказано, більшість моделей асоціативної пам'яті базуються на хебівському принципі навчання [67], який стверджує, що вага зв'язку між двома нейронами збільшується, "якщо один повторювано бере участь в активації іншого" (те, що це вже втретє в роботі дається визначення, свідчить про його важливість). А також використовується лінійна сумація вхідних імпульсів. Проте, за останні 15 років зібрали достатньо свідчень, що нейрон виконує значно складніші обчислення зі значно складнішими правилами навчання [104, 162, 36]. Дендритне дерево більше не може розглядається як пасивний приймач, воно є активним елементом обробки інформації [22, 155, 84, 116]. Крім того, теоретичні результати показують, що нелінійні дендритні властивості збільшують ємність пам'яті нейронної мережі [138]. Нещодавно було відкрито, що нові синапси утворюються на дендриті локально утворюючи кластер [46], тим самим збільшуючи супралінійне підсумовування і працюючи як детектор збігу. Все більше і більше даних підтримують ідею про те, що ми повинні знайти нові способи навчання штучних нейронних мереж [13, 23, 167].

На рис.2.2а зображено типовий пірамідальний нейрон в корі головного мозку. Як видно, дендритні та аксонні відростки можуть розростатися до сотень мікронів, при розмірі тіла клітини (соми) в середньому 20 мкм. Довгий час залишалося таємницею, яка роль зв'язків, що розташовані далеко від соми, так як активність від них швидко затухає і майже ніякого ефекту на стан клітини немає. Пізніше показали, якщо в локальному дендритному сегменті напруга більша певного порогу, то відбувається супралінійна сумація [91, 104, 155, 98, 22]. На рис.2.2в схематично показано результати активації нейрона від стимуляції двох

різних та двох близьких місцях. По осі y , відмічено виміряну активацію, або як її ще називають збуджувальний постсинаптичний потенціал (EPSP). По осі x , показано очікувану активацію у випадку лінійної сумачії вхідних стимулів. Як видно, якщо активують два різних дендрита, то нелінійні ефекти відсутні (показано зеленими квадратами), так як в типовій моделі нейрона. Однак, якщо активуються сусідні місця на одному дендриті, то активація сумується нелінійно і нагадує функцію порогу або сигмоїду. Такий ефект називають дендритний спайк, який забезпечується наявністю потенціал залежних NMDA, Ca^{2+} або натрієвих каналів. Дендритний спайк відрізняється від звичайного в аксоні, він не є самопідтримуваним, тобто з відстанню згасає, проте локально має схожі характеристики – при перетині певного граничного значення напруги на мембрані, відбувається швидке наростання активації. Такі дендритні обчислення змінюють парадигму збереження пам'яті в біологічних нейронних мережах [84, 23, 85].

Дендритний спайк спостерігається в лабораторних умовах, проте чи має місце цей ефект в живій діючій мережі? Раніше, вже було показано, що моделі нейронів, що включають збіги локально активованих синапсів на дистальних дендритах, краще відтворюють дані реальних нейронів з кори макак [157]. В нещодавній роботі 2012 року [164] показали, що дендритні сегменти активуються локально також і в робочій мережі. На рис. 2.3а червоним кольором показано більшу активність активації дендритів. Як видно, вона зосереджено локально, а отже спостерігаються нелінійні властивості дендритів. Зовсім недавно, в 2018 році, в роботі [46] показали, що при навчанні, нові синапси від активних нейронів утворюються на локальних дендритних сегментах. Варто зазначити, що така кластеризація синапсів спостерігалася на місцях зникнення попередніх шипиків; це ілюструє рис. 2.3б.

Звичайно, були розроблені моделі, які враховують нелінійну сумачію на дендритних сегментах, зокрема, модель кластерону [114, 115], сігма-пай нейрон [123, 134] та більш сучасна модифікація НТМ нейрон [62]. Проте, жодна не

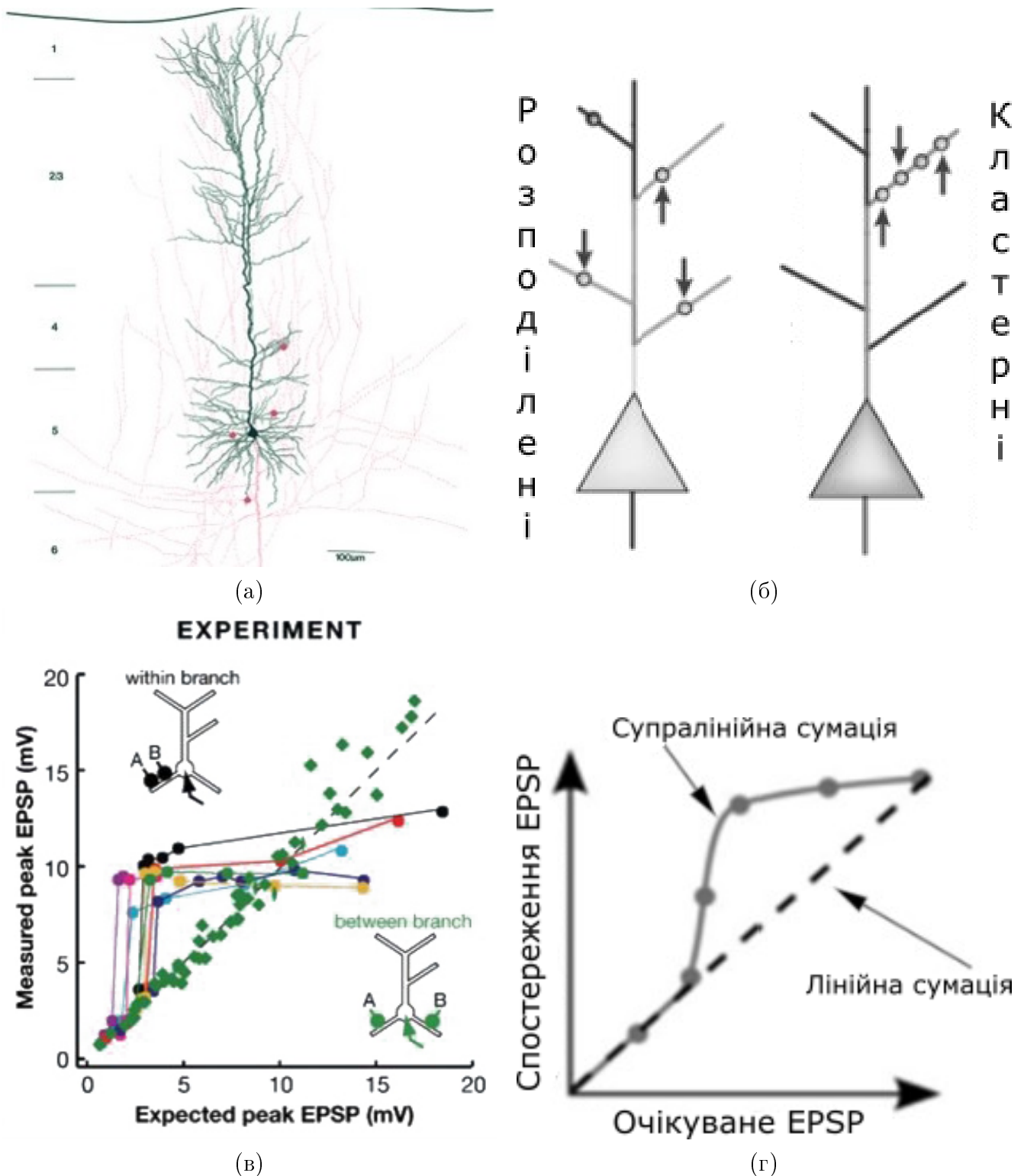


Рис. 2.2: а) Типовий пірамідальний нейрон. б) Зображення розподілених синапсів зліва і кластеризовані з правого боку. Змінено з [55] в) Результати експерименту активації різних частин дендритного дерева [104]. Спостерігається нелінійна сумація у випадку активації локального сегменту, та лінійна при активації різних віддалених місць г) Різниця між лінійним і супралінійним підсумовуванням.

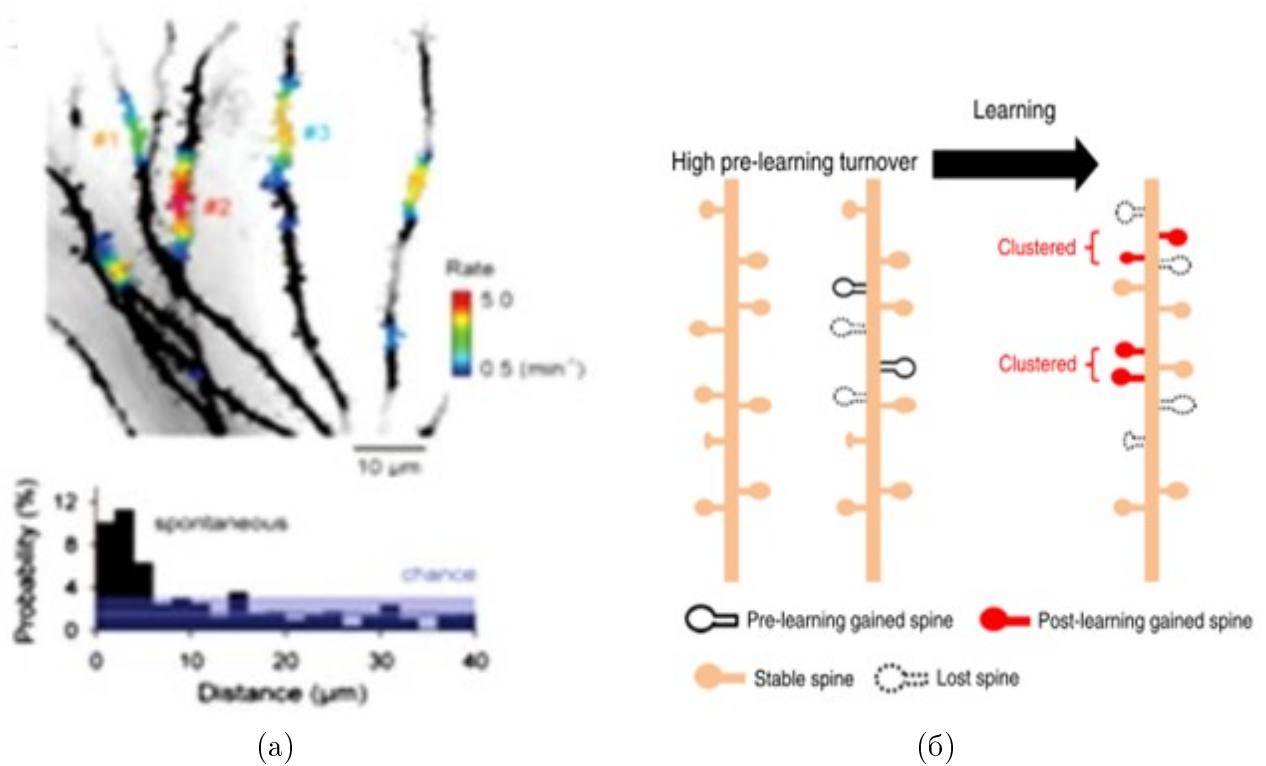


Рис. 2.3: а)Зображення активації дендрита в живій тканині (in vivo), що свідчить про локальність активації [164] б) Схематичне зображення шипиків на дендриті до і після навчання. Дані свідчать, що при навчанні нові синапси утворюються локально [46].

досліджувалася на ємність асоціативної пам'яті, що і зроблено в даній роботі.

Інтуїція зберігання кореляцій вищих порядків

Як уже не одноразово писалося, стандартний спосіб активації нейрона передбачає зважену суму вхідних сигналів і застосування деякої нелінійної функції $y = g(\sum_i w_i x_i)$. Однак, лінійне підсумовування призводить до багатьох хибно позитивних помилок. Наприклад, розглянемо мережу з 10 вхідними нейронами і одним вихідним з лінійною активацією $y = \sum w_i x_i$, та хеббівським правилом навчання $\Delta w_i = x_i y$ (рис. 2.4). На етапі навчання перший патерн активації $x_2 = 1, x_8 = 1, y = 1$ призводить до збільшення ваг w_2 і w_8 . Другий патерн $x_2 = 1, x_5 = 1, y = 1$ призводить до збільшення w_2 і w_5 . На етапі відтворення, патерн $x_5 = 1, x_8 = 1$ викликає помилкову активацію $y = 1$ тому, що w_5 і w_8 ненульові. Проте, такого патерну ніколи не було, таким чином в моделях асо-

ціативної пам'яті та нейроном з лінійною сумацією при сильно корельованих патернах виникає багато помилок.

Давайте розглянемо інший випадок, коли нейрон зберігає комбінації активних вхідних нейронів. Нехай у попередньому прикладі цей нейрон зберігає $x_2 \cdot x_5$ і $x_5 \cdot x_8$. Тоді, якщо на вході $x_5 = 1, x_8 = 1$, нейрон не активується, оскільки він не зберіг спільну активацію нейронів $x_5 \cdot x_8$. У цьому випадку ми можемо сказати, що нейрон зберігає потрібні кореляції $x_i x_j y_k$, між двома вхідними нейронами і самим собою. Тому, було б добре мати модель нейрона, яка вивчає збіги в активації нейронів, і не тільки попарно, як в оригінальному принципі Хебба.

Описаний вище ефект нелінійності активації дендритів обчислювально відповідає детектору збігів, коли саме певна комбінації нейронів здатна активувати нейрон, а не сума вхідної активації довільних нейронів, що вище порогу. Враховуючи, що середня довжина дендритного дерева типового пірамідального нейрона в корі головного мозку порядку 10^4 мкм, а дендритний спайк може виникнути на сегменті розміром порядку 10 мкм, можна говорити про множинне детектування збігів одним нейроном.

Схоже, що саме таким шляхом вирішується проблема розрізнення корельованих патернів в реальних нейронних мережах.

Модель

Раніше було враховано дендритні обчислення в модель нейрона 2.1 сігма-пай [114, 136, 62], і в даній роботі вперше було побудовано та досліджено асоціативну пам'ять з таким нейроном. Для цього було використано спрощений алгоритм

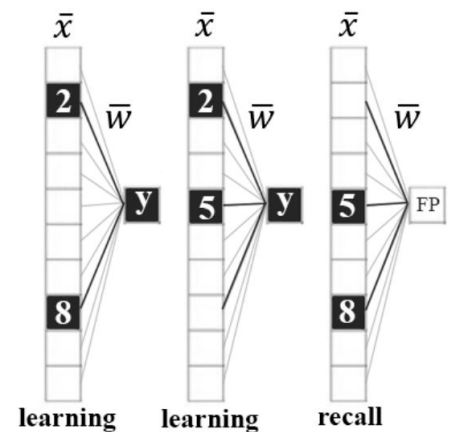


Рис. 2.4: Лінійне зважене підсумовування і простий хеббівський принцип приводять до хибної активації нейрона.

навчання 2.2.

$$y_j = \theta\left(\sum_{m=1}^{m_j} \prod_{i \in G_m^j} x_i\right) \quad (2.1)$$

$$\forall j : y_j = 1, \begin{cases} G_{m_j+1}^j = \{i | i \in A_x^k\} \\ m_j = m_j + 1 \end{cases} \quad (2.2)$$

Де $\theta(x) = \{1, x > 0; \text{інакше}, 0\}$ - крокова функція,

$$G_{m_j}^j$$

- дендритний сегмент, або кластер, j -го нейрону. При навчанні для кожного активного нейрону у векторі \mathbf{y} вибираються k індексів активних вхідних нейронів і записується в новостворений кластер. Таким чином якщо всі k вхідних нейронів, що записані в кластер активні $x_i = 1$, тоді добуток ненульовий і вихідний нейрон активується $y_j = 1$. Схематично модель нейрона та схема навчання зображена на рис.2.5а.

Рівняння 2.1 описує функцію активації на етапі відтворення вихідного вектору. Нейрон має m кластерів, кожен з них містить k індексів вхідних нейронів $G_m^j = i_1, i_2, \dots, i_k$. Якщо вхідний патерн перекривається з будь-яким з кластерів, тоді нейрон активується. Важливо, що кожен кластер це не окремий дендрит, оскільки розмір кластера в даній моделі може бути малий $k < 5$. Біологічно один дендрит може містити багато кластерів, але в моделі зроблене штучне розділення для дослідження обчислювальних властивостей.

Рівняння 2.2 описує створення кластерів на етапі навчання. Кожний активний нейрон у векторі \mathbf{y} виділяє k випадкових активних нейронів у вхідному векторі \mathbf{x} та зберігає у новий кластер $G_{m_j+1}^j$. Кількість кластерів не фіксується, і вона може динамічно зростати для нових патернів. Модель має вільний параметр k - розмір кластера. Випадок $k = 2$ відповідає попередньому прикладу з навчанням потрійних кореляцій.

Модель отримала назву розріджений кластерон так як натхненна оригінальною моделлю кластерона, представленою в статті Б.Мела [114]. Модель, заснована на тих же принципах виявлення збігу у активації, однак, змінено правило навчання, що дозволяє більш легко оцінити ємність пам'яті. Крім того, для активації вихідного нейрона має відбутися повний збіг принаймні одно з кластерів з вхідним патерном. Таким чином, розмір кластера може бути невеликим, не більше п'яти синапсів, що достатньо для відстеження більш високого порядку кореляції в розріджених вхідних векторах. Модель намагається включити сучасні погляди на дендритні обчислення [22, 168, 84] і як вони впливають на навчання нейронів. Також було використано результати, що показують важливість розрідженої активності в різних областях мозку: в гіпокампі [125, 140], деяких частинах базальних гангліїв і особливо в поверхневих шарах кори головного мозку [106, 131]. Однак модель використовує мінімальні біологічні деталі задля полегшення теоретичного дослідження та можливості виведення аналітичної формули оцінки ємності пам'яті.

Формування кластерів у представленій моделі може інтерпретуватися як ідеальна форма структурної пластичності [41]. Структурна пластичність не змінює силу зв'язку, а змінює місце цього зв'язку (до якого нейрона приєднаний і де саме на дендритному дереві). Розглянемо типовий пірамідний нейрон. В його околі знаходяться десятки тисяч аксонів від інших нейронів. Проте лише невелика частина аксонів з'єднані, тоді як інші лише потенційно можуть приєднатися. Нейрон може утворювати багато з'єднань з одним і тим самим аксоном в різних місцях на дендритному дереві. Припустимо, що з усіх активних вхідних нейронів є принаймні два в яких аксони розташовані близько (Рис. 2.56). Синапси з цих корельованих вхідних аксонів, перетворюються з потенційних в фактичних і приєднуються локально на дендритний сегмент та формують кластер. Останні експерименти частково підтверджують цю ідею [46]. Таким чином, завдяки структурній пластичності, аксони постійно намагаються створити стабільні зв'язки шляхом пошуку нових місць. Події в середовищі, кодуються в

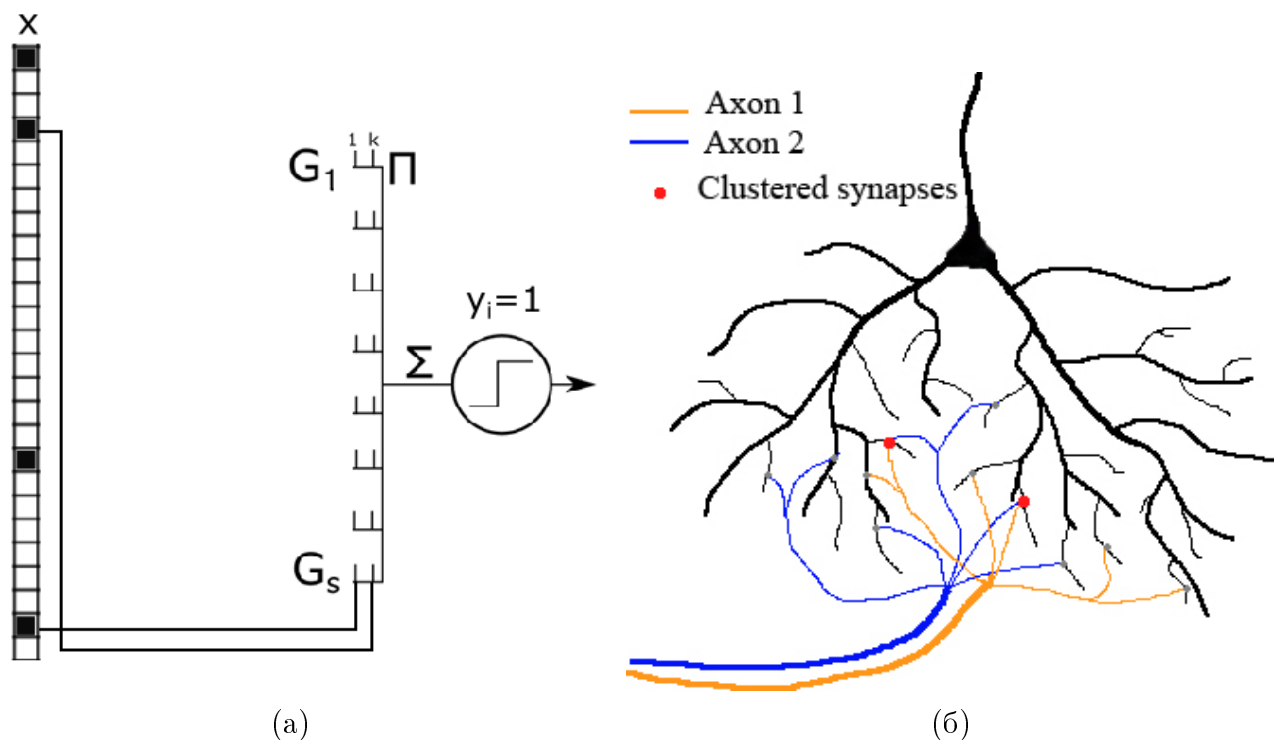


Рис. 2.5: а) Модель нейрона з дендритами розріджений кластерон. Нейрон має S кластерів, що зберігають k активних вхідних нейронів, що активували цю даний нейрон. Тільки повторна активність саме цієї групи з k нейронів може активувати кластер. б) Ілюстрація випадково розміщених аксонів. Близько розташовані аксони можуть сформувати локальні синапси, які сформують кластер на дендриті.

корельовані групи нейронів, які шляхом кластерного розміщення активних синапсів зв'язуються разом створюючи асоціативну пам'ять. Це можливий спосіб пояснити, яким чином структура світу може бути віддзеркалена в структурі зв'язків в нейронній мережі.

Ініціалізація асоціативної пам'яті

Запропонована модель застосовується для створення R асоціацій між двома нейронними популяціями $\{(\mathbf{x}^\mu \mapsto \mathbf{y}^\mu) : \mu = \overline{1, R})\}$, які представлені у вигляді бінарних векторів \mathbf{x} і \mathbf{y} розміру $|\mathbf{x}| = N_x$ і $|\mathbf{y}| = N_y$. Кількість активних нейронів a набагато менша за сукупну кількість нейронів $a \ll N$. Розрідженість визначається як $s = \frac{a}{N}$, тому в даному випадку розрідженість висока $s \ll 1$.

На етапі навчання всі R пари асоціюються за рівнянням. 2.2 і для кожного

нейрона в \mathbf{y} створюється набір кластерів. На етапі відтворення на вхід подається вектор \mathbf{x}^l без спотворення та відтворюється вектор $\hat{\mathbf{y}}$ відповідно до рівняння 2.1. Асоціативна пам'ять повинна отримати вектор максимально наближений до оригінального \mathbf{y}^l .

Далі, досліджено дві варіації моделі, одноетапна та багатоетапна. В першій варіації вихідний вектор відтворюється за одне застосування рівняння 2.1. В другій – створюються автоасоціативні зв'язки ($\mathbf{y} \mapsto \mathbf{y}$) і використовується така сама модель, лише вектор \mathbf{x} змінюється на \mathbf{y} . Для цього випадку вхідний вектор відтворює вектор $\mathbf{x}^l \rightarrow \hat{\mathbf{y}}_1$, тоді $\hat{\mathbf{y}}_1 \rightarrow \hat{\mathbf{y}}_2$ і після декількох кроків рівняння 2.1 отриманий вектор сходиться до \mathbf{y}^l .

Якість пам'яті залежить від критерію допустимості ϵ . Для одноетапної асоціативної пам'яті, критерій вибирається таким чином, щоб не було більше ніж 1% хибно активних клітин. Це еквівалентно обмеженню ймовірності помилкової активації клітини, $p(\hat{y}_i = 1 | y_i = 0) \leq \epsilon$, де $\epsilon = 0.01$. Для багатоетапної пам'яті для перших отриманих векторів можна послабити умову максимальної кількості помилок, так як з кроками автоасоціації вони зменшуватимуться. Тому, перший отриманий вектор $\hat{\mathbf{y}}^l$ може мати більше ненульових елементів, ніж оригінальний вектор \mathbf{y}^l . Однак, можлива ситуація, коли вектор буде сходиться до іншого вектора $\mathbf{y}^f, f \neq l$. Ми використовуємо таку саму величину критерію $\epsilon = 0.01$ для обмеження кількості неправильно відтворених векторів з поміж R , і обмежуємо відповідну ймовірність $p(\mathbf{y}^f | \mathbf{x}^l, f \neq l) \leq \epsilon$. Встановивши критерій допустимості, ми можемо теоретично оцінити ємність пам'яті та вивести наближену аналітичну формулу.

Далі будуть досліджені наступні основні параметри будь-якої асоціативної пам'яті:

1) Максимальна збережена інформація. Визначається як відношення максимальної взаємної інформації між відтвореним та записаним вихідними векторами до кількості затрачених ресурсів. Тобто, скільки можна записати пар векторів $\{\mathbf{x}^\mu \rightarrow \mathbf{y}^\mu\}$ так, щоб зберігалось правильне відтворення асоціацій, при

затраті мінімум ресурсів (параметрів).

2) Стійкість до помилок. Визначається як максимальне допустиме відхилення адресного вектору $\tilde{\mathbf{x}}$ від \mathbf{x}^μ при якому ще видобувається збережений вихідний вектор $\mathbf{y}^\mu = f(\tilde{\mathbf{x}}^\mu)$

3) Ефективності роботи. Визначається як час необхідний для обрахунку функції $f(\cdot)$.

Теоретичний аналіз

Припустимо, що під час етапу навчання R пар векторів (\mathbf{x}, \mathbf{y}) були асоційовані відповідно до рівняння 2.2 і створені кластери активних нейронів. Далі, на етапі відтворення застосовується рівняння 2.1 для всіх збережених пар. Щоб оцінити ємність пам'яті, потрібно визначити скільки нейронів у векторі \mathbf{y} стають активними, або, яка ймовірність активації.

Для отримання аналітичного виразу ймовірності, використовується трюк зі збільшенням розмірності задачі. Набагато простіше вирішити аналогічне завдання з новим розміром вхідного вектору $\dot{N}_x = \binom{N_x}{k}$ та рівнем активації $\dot{a}_x = \binom{a_x}{k}$, де $\binom{N_x}{k}$ - біноміальний коефіцієнт. Далі, у підрозділі "Результати" показано, що це аналогічне завдання правильно оцінює ймовірність для оригінальної задачі.

Введемо віртуальні бінарні ваги w розміру $[N_y, \dot{N}_x]$, і використаємо модифіковане Хеббівське правило навчання $w_{ij} = y_i(\dot{\mathbf{x}})$, де $f(\dot{\mathbf{x}})$ повертає один випадковим чином обраний ненульовий елемент з нового вхідного вектору. Таким чином, кожен нейрон \mathbf{y} з'єднується точно в одному місці в розширеному просторі за один крок навчання. Один крок навчання зменшує розрідженість матриці ваг для одного нейрона на $s_0 = \frac{1}{\dot{N}_x}$. Після R збережених пар кінцева розрідженість становить $s_w = 1 - (1 - s_0)^{R s_y}$, де враховано частоту активації нейронів в \mathbf{y} .

Розглянемо випадок, коли вектор \mathbf{x} не спотворений. Таким чином він відтворює всі клітини з вихідного вектору та додатково активує інші клітини, які

і називаються хибно позитивні помилки. Ймовірність того, що клітина в \mathbf{y} стає хибно активною, дається формулою: Eq. 2.3, з коригуванням розрідженості s_y .

$$p(\hat{y}_i = 1 | y_i = 0) \equiv p = (1 - (1 - s_w)^{a_x})(1 - s_y) \quad (2.3)$$

Використовуючи критерій допустимості $\epsilon = 0.01$ для одноетапної пам'яті, умову $p \leq \epsilon$, і граничний випадок $s_0 \rightarrow 0$, отримуємо максимальну кількість збережених пар відносно заданого критерію:

$$R_{\max} = \frac{\epsilon}{s_y \acute{s}(1 - s_y)} \quad (2.4)$$

, де $\acute{s} = \frac{a_x}{N_x}$.

Ймовірність помилки для багатоетапної асоціативної пам'яті задана формулою 2.5. В даному випадку, припускається, що автоасоціативна пам'ять ідеальна, тобто працює як пошук найближчого сусіда. Перший відтворений вектор $\hat{\mathbf{y}}_1$ порівнюється зі всіма збереженими, і видобувається найближчий. В якості відстані вибрана відстань Хеммінга.

$$p(\mathbf{y}^f | \mathbf{x}^l, f \neq l) \equiv p_2 = 1 - (1 - p^a)^R \quad (2.5)$$

Використовуючи $p_2 \leq \epsilon$ та граничний випадок $s_0 \rightarrow 0$, отримуємо:

$$R_{\max} = \left(\frac{\epsilon}{(s_y \acute{s}(1 - s_y))^a} \right)^{\frac{1}{a+1}} \quad (2.6)$$

Таким чином, основний результат приведенного теоретичного аналізу показує, що максимальна кількість збережених патернів для асоціативної пам'яті з дендритними нейронами залежить від розрідженості вихідного вектору та розрідженості вхідного в комбінаторно розширеному просторі. Міру розширення визначає розмір кластера на дендритному сегменті, що еквівалентно порядку кореляції.

Чутливість до шуму та швидкість відтворення

Попередній аналіз показав теоретичну оцінку ємності пам'яті у випадку без шуму, коли вхідний вектор дорівнює одному із збережених $\hat{\mathbf{x}}^l = \mathbf{x}^l$. У загальному випадку з наявністю шуму, вхідний патерн має λ правильно і β неправильно активних клітин, так що $\hat{a}_x = \lambda + \beta$. Для випадку $\hat{a}_x = a_x$, рівняння (2.3) залишиться без змін, оскільки число активних нейронів залишається незмінним. Однак, тепер не всі клітини з правильного вихідного вектору активні і це потрібно враховувати у помилково позитивній (false positive) ймовірності:

$$p(\hat{y}_i = 0 | y_i = 1) \equiv p_{fp} = \frac{\binom{a_x - \lambda}{k}}{\binom{\lambda}{k}} \quad (2.7)$$

Сукупна помилкова ймовірність включає в себе як помилково позитивні, так і помилкові негативні (false negative) активні клітини. З умови $p + p_{fp} \leq \epsilon$ можемо аналогічно оцінити максимальну кількість збережених пар патернів R_{max} .

У випадку відсутності неправильної активації $\beta = 0$ ми маємо лише частково активний вхідний вектор $\lambda = \hat{a}_x < a_x$. Це зменшує помилкові негативно активовані клітини в рівнянні 2.3 шляхом зміни $\hat{a}_x = \binom{\lambda}{k}$. Але це також призводить до збільшення помилково позитивно активних клітин.

Що стосується швидкості відтворення, то для отримання вектору \mathbf{y} потрібно перебрати всі N нейрони та всі k синапси з-поміж S кластерів. Таким чином, виконується - NSk операцій. Порівняно з моделлю Вілшоу, в якій $aN + N$ операцій Knoblauch2010, розріджений кластерон повільніший. Однак, якщо розглядати паралельну архітектуру подібну до мозку, де окремі дендритні сегменти виконують обчислення, час роботи є порівнянним, оскільки всі синапси працюють паралельно. Варто додати, що паралельна реалізація набагато швидша порівняно зі звичайними таблицями пошуку (look up tables), які використовуються на послідовних комп'ютерах.

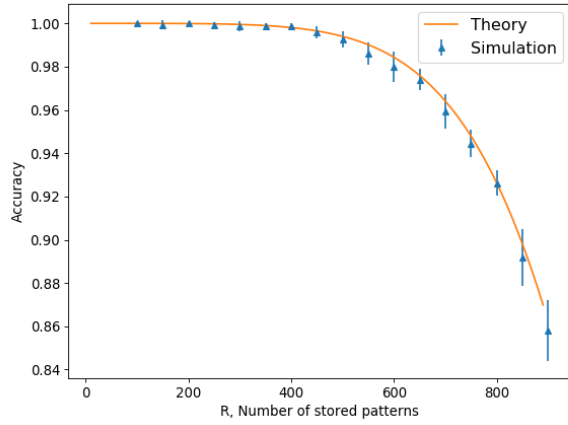
Результати

Для перевірки теоретичних розрахунків порівняно аналітичну формулу ємності пам'яті з результатами чисельних експериментів. Створено R випадкових пар патернів $(\mathbf{x}^\mu, \mathbf{y}^\mu), \mu \in \{1, \dots, R\}$ з фіксованою кількістю активних нейронів $a_x = a_y = 7$, з розміром мережі $N_x = N_y = 100$ та розміром кластера $k = 2$. Теоретична точність розраховується відповідно до $Ac^{theory}(R) = 1 - p_2$ для багатоетапної пам'яті та порівнюється з результатами моделювання $Ac^{sim}(R) = \frac{\#error\ retrieval}{R}$ для різних значень R . Оскільки вектори є випадковими, проведено кілька експериментів і усереднено результати. Як видно з (рис. 2.6а), теоретичний розрахунок узгоджується з результатами моделювання.

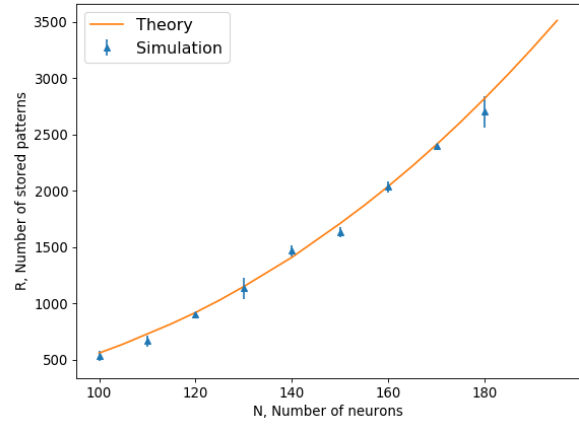
Для наступного експерименту проводиться порівняння максимальної кількості збережених патернів R_{max}^{theory} та R_{max}^{sim} при різних розмірах мережі N . Як видно з рис. 2.6б теоретичний результат повторює результат симуляції. Таким чином, підтверджено правильність розв'язку задачі в розширеному просторі та її еквівалентність до поставленої задачі розрідженого кластерона. Це дає цікавий висновок, що через відстеження кореляцій вищих порядків нейрони розширюють розмірність вхідного простору в якому легше розрізнити різні патерни.

Залежність ємності від параметрів моделі

Основними параметрами розрідженого кластерона є k - розмір кластера або порядок кореляцій, a - кількість активних нейронів, та N - загальна кількість нейронів. На малюнку 2.7а представлено результати максимальної кількості збережених патернів з фіксованим рівнем активації $a_x = a_y = 6$ за різних значень N і k при багатоетапній пам'яті. Чим більший розмір мережі та розмір кластера, тим вища ємність пам'яті. Цікаво, що один нейрон може зберігати величезну кількість патернів: за певних параметрів це може бути порядку мільйона патернів на нейрон, хоча й це вимагатиме величезної, біологічно нереалістичної кількості синапсів. Ресурси в реальному мозку, його кількість синапсів



(a)



(б)

Рис. 2.6: Порівняння теоретичної та експериментальної точності та ємності мережі: а) Залежність точності відтворення від кількості збережених патернів з рівнем активації $a = 7$ та розміром кластера $k = 2$. Від $R = 400$ точність починає падати. Теоретична крива відмінно повторю експериментальні дані. б) Залежність максимальної кількості збережених патернів від кількості нейронів. Інші параметри залишаються незмінними. Ємність збільшується з розміром мережі. Теоретичний та експериментальні результати узгоджуються.

та загальна довжина дендритного дерева обмежені. Проте математично абстрактний нейрон не обмежений і може забезпечити значно більшу ємність мережі в цілому.

Наступний експеримент досліджує, як кількість активних нейронів впливає на ємність пам'яті. Для фіксованого розміру кластера $k = 3$ порівняно максимальну кількість збережених патернів для різних значення a і N (рис. 2.76). Результати показують, що мережа має найбільшу ємність для $a = 6$ для заданого діапазону розмірів мережі, що узгоджується зі звичайними моделями асоціативної пам'яті з лінійною сумациєю та парними кореляціями, де оптимальним рівнем активації є $a = \log_2 N$ [89]. Оптимальний рівень активації відповідає розрідженості 2 – 6% для різних розмірів мережі, що порівняно з біологічними даними розрідженості сенсорних та асоціативних областей кори головного мозку.

Одноетапну асоціативну пам'ять набагато легше реалізувати і вона вимагає менше ресурсів, однак, додавання латеральних автоасоціативних зв'язків

збільшує ємність. На рис. 2.8а представлено результати порівняння одно- та багатоетапної асоціативної пам'яті. Як видно багатоетапна реалізація має на порядок вищу ємність.

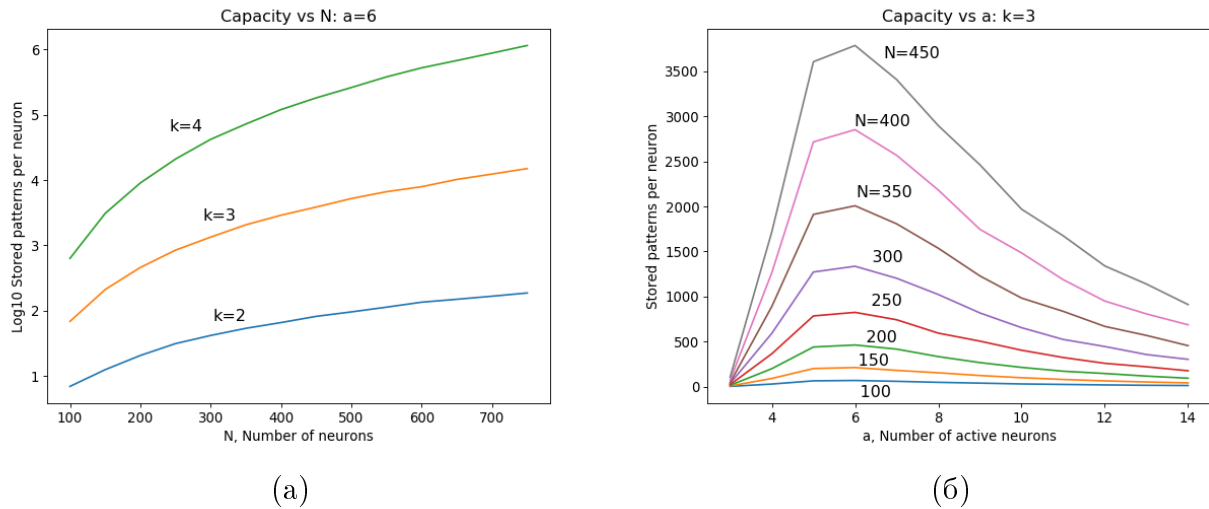
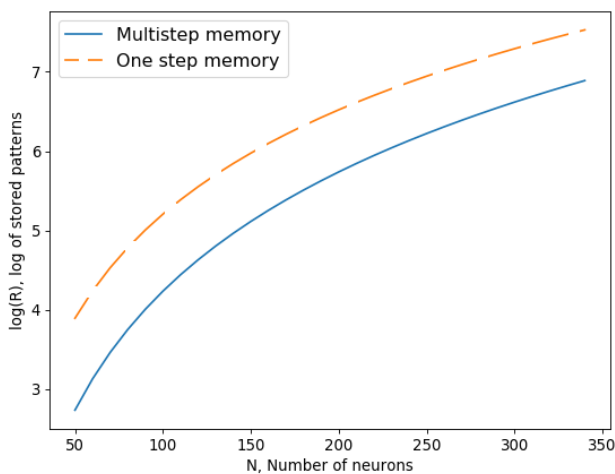


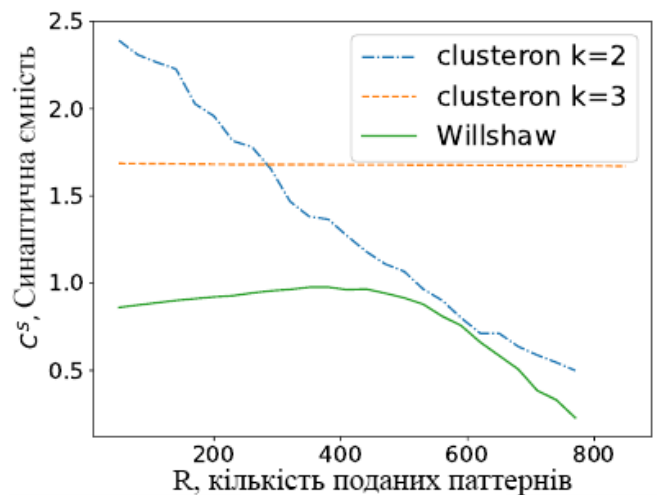
Рис. 2.7: Максимальна кількість збережених патернів для різних параметрів. а) Ємність для різних розмірів мережі та різних розмірів кластерів для фіксованого рівня активація $a = 6$. Вертикальна вісь у логарифмічній шкалі показує збережену кількість патернів поділену на кількість нейронів. Чим більше розмір мережі та кластерів, тим більша ємність. Для мережі $N = 700, k = 4$ один нейрон може зберігати мільйон патернів. б) Ємність залежно від рівня активації для різних розмірів мережі та фіксованого розміру кластера $k = 3$. Ємність вимірюється як відношення кількості успішно відтворених патернів до загальної кількості нейронів для заданого критерію допустимості $\epsilon = 0.01$. Ємність має оптимальний рівень активації $a = 6$. Чим більше мережа, тим більше патернів може зберігати один нейрон. Для $N = 400$ та $k = 3$ один нейрон може зберігати майже 3000 патернів.

Порівняння з класичною моделлю

Щоб порівняти різні моделі асоціативної пам'яті, недостатньо лише взяти до уваги максимальну кількість збережених патернів, потрібно також врахувати кількість синапсів та обсяг інформації для їх опису. В роботі [89] зазначено основні міри для ємності пам'яті:



(а)



(б)

Рис. 2.8: (а) Порівняння максимальної кількості збережених патернів для одно- та багатоетапної асоціативної пам'яті. Додавання автоасоціативних зв'язків збільшує ємність на порядок. (б) Порівняння синаптичної ємності моделі Вілшоу та розрідженого кластерона з $k = 2$ і $k = 3$ синапсами в кластері. Синаптична ємність моделі Вілшоу збільшується з кількістю збережених патернів, а потім падає через великий рівень помилок. Синаптична ємність для $k = 3$ на даному діапазоні кількості патернів постійна, оскільки пам'ять не насичена для якої $R_{max} > 800$. Синаптична ємність для $k = 2$ зменшується, але вона більша, порівняно з моделлю Вілшоу.

Мережева ємність:

$$C = \frac{T(N, a, R_{max})}{\#contacts} [bit/contact]$$

де $T(N, a, R) = RNI(\frac{a}{N})$ ємність каналу, що можна інтерпретувати як максимальна збережена інформація. R_{max} - максимальна кількість збережених пар асоціацій. $I(p) = p \log_2(\frac{1}{p}) + (1 - p) \log_2(\frac{1}{1-p})$.

Синаптична ємність:

$$C^S = \frac{T(N, a, R_{max})}{\#nonsilent synapses} [bit/synapse]$$

. Ця міра подібна до попередньої, але враховує тільки ті зв'язки в яких ненульові ваги, тобто ті синапси, які реально використовуються. Ця міра є найбільш важливою для практичної реалізації моделей пам'яті в hardware, так як там важливо фізична наявність чи відсутність зв'язку.

Інформаційна ємність:

$$C^I = \frac{T(N, a, R_{max})}{\#bits \text{ of required physical memory}}$$

Вона найбільш точно характеризує ємність пам'яті, так як враховує не тільки кількість зв'язків, але й кількість інформації потрібної для опису кожного зв'язку. Наприклад, в моделі Вілшоу зв'язки бінарні, тому потрібно 1 біт на зв'язок.

В таблиці 2.1, наведені формули для обрахунку різних мір для розрідженого кластерона. Для умови $a \ll N$ можна обрахувати наближення $T(N, a, R_{max}) = R_{max}NI(\frac{a}{N}) \approx R_{max}a \log_2(\frac{a}{N})$. Також, враховуючи, що в моделі нейрона з дендритами за R асоціацій утворюється Rka зв'язків, звідси і отримуємо значення різних мір ємності для запропонованої моделі.

Далі приводиться порівняння синаптичної ємності для розрідженого кластерона та модифікованої моделі Вілшоу. Щоб провести справедливе порівнян-

Назва міри	Формула означення	Ємність моделі
Мережева ємність	$C^N = \frac{T(N,a,R_{\max})}{\text{зв'язків}} [bit/contact]$	$C^N = \frac{1}{k} \log_2(\frac{N}{a})$
Синаптична ємність	$C^S = \frac{T(N,a,R_{\max})}{\text{ненульових зв'язків}} [bit/synapse]$	$C^S = \frac{1}{k} \log_2(\frac{N}{a})$
Інформаційна ємність	$C^I = \frac{T(N,a,R_{\max})}{\text{інформація на зв'язок}}$	$C^I = \frac{1}{k} \log_2(\frac{N}{a \log_2(N)})$

Табл. 2.1: Міри ємності та формули для обрахунку для запропонованої моделі.

ня, використано одноетапну реалізацію. Також, порівнюється саме синаптична ємність так як в кластероні немає нульових зв'язків (як у моделі Вілшоу). Модифікації моделі Вілшоу включають зміну порогового значення $\theta = a$ на процедуру $kWTA()$, яка працює краще [56]. А також, враховуються тільки ненульові зв'язки в синаптичній ємності.

З обчислювальних експериментів отримано залежність $C^S(r)$, де r - це число представлених патернів. Для моделі Вілшоу, представлення нових патернів збільшує кількість використаних синапсів, і після певного порогу збільшується рівень помилок, і ємність падає внаслідок заповненості матриці ваг (рис. 2.86. Для розрідженого кластерона з $k = 3$ синаптична ємність є константою, оскільки $R_{max} > 800$, що відповідає теоретичному передбаченню. При $k = 2$ синаптична ємність падає, однак вона вища порівняно з моделлю Вілшоу.

Так як в моделі Вілшоу багато нульових зв'язків, мережева ємність C^N значно менша ніж в кластерона.

У моделі кластерона кожен кластер зберігає індекси вхідних нейронів, тому потрібно $\log_2(N_x)$ біт інформації для запису кожного індексу. Проте, в моделі Вілшоу використовуються бінарні синапси, тому вона має вищу інформаційну ємність C^I , ніж у кластерона. Тим не менш, у випадку практичної реалізації моделей пам'яті в hardware, синаптична ємність важливіша, так як визначає необхідну кількість фізичних контактів.

Слід зазначити, що дві моделі мають один і той самий порядок для різних ємностей зберігання, однак, розріджений кластерон має набагато вищу максимальну кількість збережених патернів. Також, безсумнівною перевагою запропонованої моделі є її здатність підлаштовуватися під вимоги пам'яті змінюючи

параметр k . Так, якщо ємність може бути не великою, можна взяти малі значення розміру кластера і зберегти ресурси. Якщо розрідженість векторів мала, тобто є багато одиниць, можна вибрати розмір кластера великим, і пам'ять буде справлятися з цільними векторами, завдяки використанню багатьох синапсів.

Обговорення

Біологічний нейрон в корі головного мозку має близько 10 000 зв'язків, і в кожний момент, принаймні, 100 активні. Для активації нейрона достатньо 15-30 вхідних нейронів, однак важливо, де саме на дендритному дереві знаходяться ці зв'язки. Якщо ці зв'язки синхронно активуються і знаходяться локально, тоді з більшою ймовірністю активується вихідний нейрон. Таким чином, окремий нейрон може зберігати величезну кількість патернів, наприклад, якщо дві мережі розміром по 700 нейронів, і відслідковується збіг трьох нейронів, то один нейрон може розпізнати до 10 тисяч різних патернів.

Оригінальність розрідженого кластерона полягає в тому, що показано як біологічне явище кластерного формування синапсів приводить до кращої асоціативної пам'яті для якої вдалося вивести аналітичні вирази ємності. Головним результатом є те, що показано, що відстеження кореляцій вищих порядків надзвичайно збільшує кількість патернів, які може надійно розпізнати нейрон. Причина цього полягає в тому, що виявлення збігів розширює розмірність вхідної мережі, в якій легше розділити патерни, і тому збільшується ємність пам'яті. Інша причина походить від розрідженої активації, що робить можливим використовувати невеликий розмір кластера. В цьому випадку при зв'язуванні з кількома клітинами з популяції, нейрон може розпізнати всю популяцію.

Раніше була спроба показати, що нейрон з нелінійними дендритами має вищу ємність [138], завдяки більшій кількості можливих комбінацій з'єднань. В даній роботі дослідження набули подальшого розвитку: побудована асоціативна пам'ять і оцінено ємність мережі в цілому, а не одного нейрону. Показано,

що розріджена модель кластерону має меншу інформаційну ємність порівняно з моделлю Вілщоу, але більш високу мережеву та синаптичну ємність, що є перевагою для апаратної реалізації. Крім того, розріджена модель кластерону має набагато більшу кількість збережених патернів і може досягати мільйонів на нейрон. Для досягнення такої ємності нейрон створює кілька з'єднань з вхідними нейронами, і активується тільки якщо вони всі активні. Це призводить до того, що два нейрони з'єднуються багатьма зв'язками в багатьох місцях, що подібно до біологічних нейронів з 10-20 контактами в середньому. Отже, не тільки сила зв'язку має значення, але й його розташування, яке зберігає додаткову інформацію.

Для моделей нейронів з індивідуальними дендритами або, як у нашому випадку кластерів, наступне питання зазвичай виникає, "чому б не використати 10 простих нейронів, замість одного складного з 10-ма дендритами?". Ці два випадки математично подібні, але все таки відрізняються. Якщо складний нейрон з'єднується з мережею розміром N складних нейронів в яких по d кластерів, то матриця зв'язків має розмірність $1 \times N \times d$. З точки зору біології, набагато ефективніше мати один аксон і генерувати один потенціал дії для передачі сигналу замість того, щоб мати десять аксонів і посилати десять потенціалів дії. Це ж стосується і апаратної реалізації штучних нейронних мереж. Отже, використання складних нейронів економить затрати енергії для генерації спайків.

Варто згадати мережі вищих порядків, які розроблялися в кінці 80-х років [9, 53, 97, 165]. Вони розширили мережа Хопфілда для врахування кореляцій вищого порядку, як в моделі кластерону. Однак вони зробили це в загальному випадку для повнозв'язної мережі, не використовуючи умову розрідженості, створивши таким чином величезні багатовимірні матриці, і є обчислювальними занадто складними для використання на практиці. Натомість, в запропонованій моделі використано ідею, що не обов'язково створювати всі зв'язки, достатньо з'єднатися з невеликою кількістю нейронів, щоб зв'язатися з популяцією в цілому [5, 6]. Це дозволяє знизити обчислювальну складність при збереженні високої

ємності пам'яті.

Існує ще одна нещодавня модель НТМ [62] з дендритами і використанням розріджених векторів [82, 83, 135, 144]. Ця модель була розроблена для зберігання послідовностей, де кожен дендрит працює як детектор збігу, але вона не застосовувалася для задачі асоціативної пам'яті. Головна відмінність порівняно з цією моделлю полягає в тому, що в даній роботі показано, що можливо використовувати кластери малого розміру для відстеження збігів і виведено аналітичні вирази для ємності пам'яті.

З обережним оптимізмом здається, що деякі головоломки, нарешті, стають зрозумілими: біологічні нейронні мережі еволюціонували шляхом збільшення дендритних дерев і утворенню синаптичних кластерів для збільшення ємності та стійкості пам'яті.

2.3 Кодування зображення в розріджені представлення

Як вже було встановлено, дендритні обчислення мають важливе значення в роботі нейрона. Далі ставиться питання, яку можливу роль вони відіграють в кодування інформації? Біологічні дані свідчать, що нейрони з таламусу створюють синапси близько до соми нейрона в шарах L4 та L6. І тому, ці синапси досить сильні, що зменшує необхідність в кооперації, як це було у випадку асоціативної пам'яті з синапсами на дальніх частинах дендритного дерева. Тим не менш, важливо дослідити питання як вплине на кодування зміна класичної активаційної функції на детектор збігів.

Стандартний спосіб кодувати щось передбачає створення словника, що ставить у відповідність об'єкт з його двійковим кодом. Наприклад, літера "А" кодується як 1000001 відповідно до ASCII. Такий підхід сформував сучасну комп'ютерну техніку, проте для біологічних організмів не має когось хто створить словник,

як людина створила для комп'ютера. Тому, вони створюють внутрішнє представлення зовнішнього середовища самі завдяки самоорганізації нейронних мереж. Якщо ми хочемо створити системи зі штучним інтелектом, ми повинні відмовитися від створення словників людьми та забезпечити можливість самоорганізації коду. Проте залишається відкритим питання, як перетворити сирі дані рецепторів в такі бінарні вектори [144].

Останнім часом в цьому напрямку було досягнуто певного прогресу. Глибокі нейронні мережі вивчають риси з "сирих" сенсорних даних, в той час як раніше ці риси задавалися вручну [95]. Однак, по суті, ці системи також потребують людей, які дадуть розмічені дані з мітками, на відміну від біологічних нейронних мереж, які використовують суміш навчання без учителя та навчання з підкріпленням.

Інший вагомий результат показує перспективи розрідженого кодування. Якщо задати вимогу реконструювати вхідне зображення через лінійну комбінацію малої кількості словників з-поміж всіх вивчених, алгоритм вивчає риси, схожі на рецептивні поля простих клітин у візуальній корі [43] (напрявлені лінії). І хоча цей підхід подібний до біологічного, проте він використовує частотний підхід кодування даних (коефіцієнти в лінійній активації – дійсні числа). Як зазначалося в першому розділі, більшість схиляються до думки, що кодування не може бути частотним для розпізнавання образів в корі головного мозку.

Також, формування розрідженого представлення було отримано в роботах Фолдіака [45] та командою Нумента [31]. Вони використали гомеостатичні принципи навчання з латеральною інгібіцією, чим змогли відтворити важливі біологічні деталі: отриманий код розріджений, для подібних стимулів код теж подібний і для більш частіших стимулів код більш розрідженіший. Останнє важливо, так як тоді статистично частіші риси кодуються меншою кількістю нейронів, що оптимізує енерговитрати організму на генерацію електричних імпульсів. Тим не менш, ці моделі обчислювально затратні і не можуть бути використані в практичних цілях. Це пов'язано і з часом навчання і з обмеженою кількістю

рис, що здатна вивчити мережа.

В даному підрозділі тестується проста модель, що утворює розріджене представлення зображення з використанням дендритних обчислень. Один нейрон працює як детектор збігів і реагує на багато рис, що збільшує загальну ємність мережі. Весь об'єкт кодується через розподілену активацію багатьох нейронів. Результати показують кодування зображення у розріджене представлення зі збереженням схожості для подібних зображень. Кожна клітина утворює кілька рецептивних полів і разом з іншими формує популяційне рецептивне поле, яке і представляє об'єкт.

Модель

В якості даних для кодування використовуються рукописні цифри з бази даних MNIST. Зображення бінаризується і отриманий вектор x розміром $28 * 28 = 784$ подається на шар нейронної мережі y . Активація розраховується відповідно до:

$$y_i = k W T A((1 - \alpha) \sum_j w_{ij} x_j + \alpha \sum_k \prod_j c_{ijk} x_j) \quad (2.8)$$

kWTA повертає k найбільш активних клітин одиницями, а інші зануляє. Активація складається з двох доданків. Перший – лінійна сумація вхідного вектору зваженого на бінарний випадковий вектор w з біноміального розподілу. Ваги створені для локальної частини зображення зі збереженням топографічних співвідношень 2.9. Другий – детектор збігів, як добуток елементів вхідного вектора з врахуванням рецептивного поля c_{ijk} . Параметр α визначає вплив лінійного члена та детектора збігів на стан нейрона. На початку коли

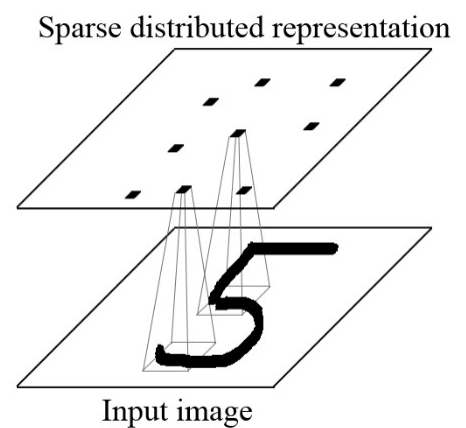


Рис. 2.9: Зображення топографічних зв'язків одного нейрона.

кластери c_{ijk} не навчені, домінувати буде перший доданок, але після навчання другий. Використовується хеббоподібне правило навчання:

$$p(c_{ijk} = 1) = \beta x_i y_k \quad (2.9)$$

яке визначає ймовірність утворення кластера, де $\beta = 0.01$.

Перетворивши кожне зображення в бінарне представлення можна застосувати різні метрики відстані для порівняння схожості кодувань.

Результати

Так як кодування мають однакову кількість активних клітин, схожість визначалася як перетин двох кодувань $\sum x^i \cdot x^j$, де x^i вектор i -го зображення. Матрицю перетину для кодувань зображень з чотирьох класів зображено на рис. 2.10 а). Чим темніше, тим вищий перетин. Цифри з того самого класу мають більший перетин, однак є екземпляри, які мають перекриття більше до іншого, чужого, класу.

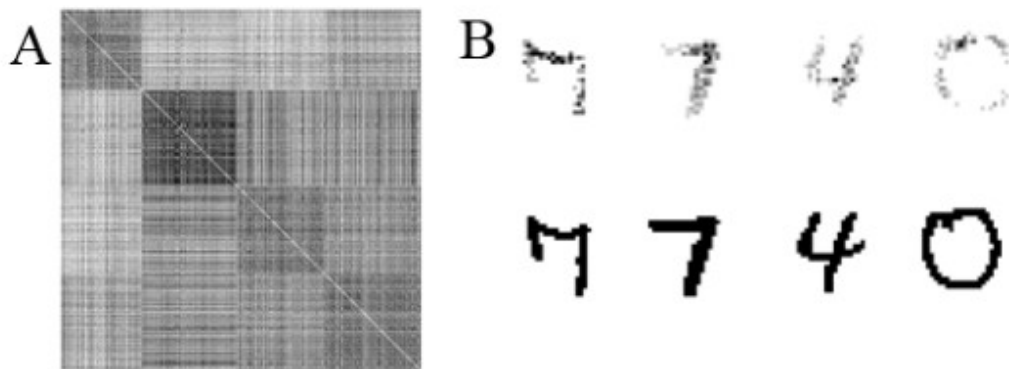


Рис. 2.10: А) Матриця перетину кодувань зображень з чотирьох різних класів. Темний колір означає більше перекриття. Видимі чорні квадрати показують високий перетин для зображень з одного класу. Б) Нижче: бінарні зображення рукописних цифр. Верхнє: популяційне рецептивне поле мережі.

Після навчання, коли для кожної клітини сформувався певний набір кла-

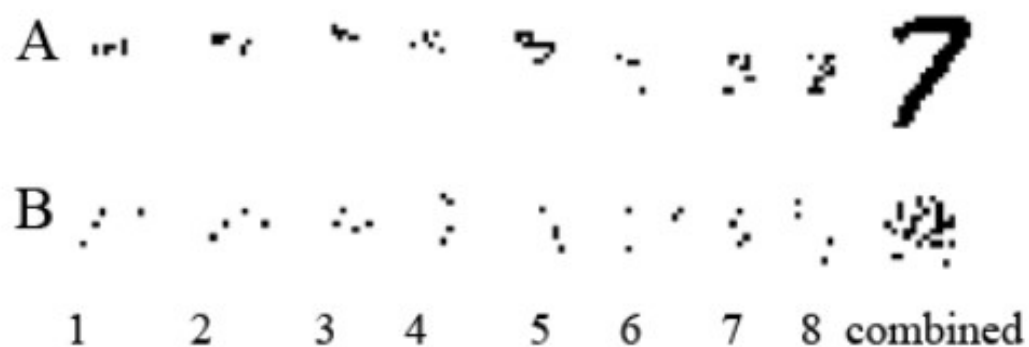


Рис. 2.11: А) Рецептивні поля активних кластерів різних нейронів, що кодують цифру. В) Рецептивне поле всіх кластерів одного нейрона та їх комбінація справа.

стерів c_{ijk} , можна візуалізувати рецептивне поле популяції. Кожний нейрон має k кластерів, де зазвичай тільки один активний. Склавши всі активні кластери для всіх активних клітин в y отримуємо зображення стимулу на рис. 2.10б), який активував дану популяцію. Внизу показано вхідне зображення, зверху – рецептивне поле популяції нейронів. Як видно, рецептивне поле досить непогано відтворює вхідний стимул. Цікаво, що жодна клітина не містить всієї інформації про вхідне зображення, але склавши всі рецептивні поля клітин, отримується стимул. Це наглядний приклад популяційного кодування. Якщо розбити популяційне кодування на окремі клітини, то видно, що кожна клітина кодує лише частину вхідного зображення, певні риси. На рис. 2.11а) показано активні кластери для деяких вибраних активних клітин. По одинці вони кодують лише певні риси, разом кодують все зображення.

Далі, виникає питання, а яке ж рецептивне поле одного нейрона? Просумувавши всі кластери c_{ijk} для j -го нейрона, отримуємо візуалізацію. На рис. 2.11б) показаний набір рецептивних полів одного нейрона, його кластерів, збіг яких він відслідковує. В сукупності, рецептивне поле зовсім не селективне. Це подібно до ідеї грубого кодування [68, 40], коли рецептивне поле одного нейрона має дуже широку селективність, але при кодуванні багатьох клітин, можна виділити потрібний стимул.

Обговорення

Цей підрозділ присвячений проблемі кодування інформації в розріджене бінарне представлення. Запропонована ідея полягає у використанні розширеної моделі нейрона, яка включає в себе дендритні обчислення. Дендрити виконують детектування збігу у вхідних даних і дозволяють одному нейрону вивчати різні рецептивні поля. Дійсно, експериментальні дані свідчать, що "складні клітини" (за термінологією Хюбеля та Візеля) формують інваріантність через різні рецептивні поля на окремих дендритних гілках [80].

В представленій моделі показано як ці рецептивні поля формуються та як їх перетин кодує все зображення через популяційне кодування. Вводиться поняття популяційне рецептивне поле, як той набір стимулів, що активує певний набір клітин. Кожна поодинокі клітина містить лише частину рецептивного поля. Цікаво, що окрема клітина має широку селективність, але на рівні популяції можливо виділити стимул. Це подібно до раніше запропонованої ідеї грубого кодування, як механізму кодування в сенсорних областях кори мозку. Результати цього розділу опубліковано в роботі автора [170]

2.4 Висновки

Даний розділ присвячено просторовій асоціативній пам'яті, в якому запропоновано нову модель пам'яті, на основі моделі нейрона сігма-пай. Ця модель включає дендритні обчислення, а саме детектування збігів. Таке відслідковування кореляцій вищих порядків дає високу ємність асоціативної пам'яті. Це поліпшення є наслідком двох ідей. Перша, – для розріджено активної нейронної мережі достатньо зв'язатися до невеликої підпопуляції, щоб з'єднатися зі всією популяцією, що кодує певну інформацію. Друга, – детекція збігів збільшує розмірність вхідних даних, тим самим збільшує роздільну здатність нейрона. Цікаво, що скоріш за все, такі обчислювальні механізми реалізовані в реальних

біологічних нейронах через синаптичну кластеризацію на дендритних деревах та нелінійні властивості потенціал залежних іонних каналів.

Модель нейрона з детекцією збігів та навчання кластерів була застосовна для кодування зображень в розріджене бінарне кодування. Врахування дендритів в нейроні призводить до формування багатьох рецептивних полів в одному нейроні. Внаслідок цього, декодувати стимул по активності лише одного нейрону не можливо, а потрібно врахувати активність багатьох нейронів, кожний з яких, кодує певні риси об'єкту.

Підсумовуючи:

1. Врахування дендритних дерев в моделі нейрона і використання розріджено розподілених представлень призводить до високої ємності асоціативної пам'яті. Окремий дендритний сегмент працює як детектор збігів і вивчає кореляції вищих порядків у вхідній нейронній активації.

2. Теоретичний аналіз мережі показав, що ємність пам'яті залежить від ступеня розрідженості активності нейронів. Чим більша розрідженість, тим вища ємність мережі. Також, теоретично можна інтерпретувати детектування збігів як комбінаторне розширення розмірності вхідної мережі, чим і пояснюється висока ємність.

3. Порівняння з класичною моделлю асоціативної пам'яті Вілшоу показало переваги запропонованої моделі в мережевій та синаптичній ємності, але програв в інформаційній ємності.

4. Застосування моделі нейрона для кодування зображення призводить до широкої селективності одного нейрону, та виникненню популяційного рецептивного поля.

Результати експериментальних досліджень даного розділу наведено в публікаціях [170, 127]:

1. Osaulenko V.M. Testing simple neuron models with dendrites for sparse binary image representation, Штучний інтелект, 2017, № 2, С. 101-108 (фахове видання)

2. Osaulenko, V., Girau, B., Makarenko, O. Henaff, P. Increasing Capacity of

Association Memory by Means of Synaptic Clustering. Neural Process. Lett. (2019)
pp 1–14 . doi:10.1007/s11063-019-10051-7 (Іноземне видання, SCOPUS)

Розділ 3

Часова асоціативна пам'ять

Даний розділ присвячений часовій асоціативній пам'яті. Спершу описано про просторово-часові патерни активності в біологічних нейронних мережах та їх властивості. Далі, навчання послідовностей формалізовано в три задачі, прогнозування, інтеграції та генерації, та перші дві розглядаються більш детально. Зокрема, побудовано модель прогнозування послідовності на основі моделі нейрона сігма-пай подібно до попереднього розділу. А також, пропонується модель часової інтеграції послідовності на основі алокації пам'яті в нейрони зі збільшеною здатністю до збудливості. Детальніше, що слід розуміти під часовою інтеграцією, буде написано далі.

В назві роботи вживається словосполучення "просторово-часова асоціативна пам'ять" яке не є загальноприйнятим. Воно запропоноване в даній роботі як поєднання терміну з нейронауки (spatio-temporal dynamics - просторово-часова динаміка) та асоціативної пам'яті.

3.1 Просторово-часова активація біологічних нейронних мереж

Вхідний сенсорний сигнал породжує просторово-часові патерни активності у нейронних структурах, що їх обробили, зокрема: у гіпокампі [34, 59, 15], візуальній V1 [51], звуковій A1 [154], моторній [102] корі головного. Припускається, що такі регулярні послідовності присутні у всій корі головного мозку [38, 62] та в багатьох інших підкортикальних структурах. Тривалий час в нейронауці тривали дискусії як представлені послідовності в нейронній активності. Один погляд вказував, що послідовності в нейронних мережах організовані як лан-

цюг, коли активний нейрон активує наступний, який далі активує наступний і так далі продовжуючи послідовність. Інший, що послідовності представлені ієрархічно, коли активний нейрон активує одразу багато інших в різних шарах нейронів, які між собою взаємодіють і відтворюють послідовність паралельно. На сьогодні отримано вичерпні відомості, що має місце другий, ієрархічний, варіант. Наприклад, у випадку послідовності дій, на перших шарах представлені окремі дії, в проміжних – нейрони кодують початок і закінчення підпослідовності дій, а в найвищих представлені поведінкові команди на абстрактному рівні, наприклад, дійти від дому до роботи.

Для того, щоб зрозуміти механізми представлення послідовностей, проводили багато біологічних експериментів. Далі наводяться основні властивості просторово-часової активації нейронів в мозку [34, 59], де під елементом послідовності слід розуміти просторовий патерн активності мережі в певний момент часу:

- Доповнення послідовностей, наприклад, коли послідовність АБВГД вивчена, то ініціація АБ реактивує всю послідовність далі, тобто, ВГД
- Чутливість до порядку: АБВ закодовано інакше, ніж ВБА
- Ефект зменшення та стабілізації активації мережі при повторюванні послідовності. Розріджена активність стає ще більш розрідженою і ті самі нейрони надійно будуть активні знову при появі навченої послідовності. Одна з можливих причин цього, що утворюються нові гальмувальні зв'язки, які починають ефективно вибирати кодування нейронів.
- Часова схожість, коли в послідовності АБВ елементи А та Б кодуються схожим чином. Кодуючі популяції мають великий перетин, який збільшується при повторенні.
- Асиметрія, коли в навченому патерні АБ, активність від А активує Б, але не навпаки.

- Диференціація послідовностей, коли нейронна активація для схожих послідовностей все більше відрізняється з повтореннями. Не плутати схожість окремих елементів в одній послідовності (попередній ефект). Це пов'язано з ефектом загасання, в результаті, послідовності АВС і ВСЕ матимуть стійкі та різні представлення.
- Асоціація послідовностей, які належать до одного контексту. Якщо послідовності Q W R кодуються в тому ж або аналогічному контексті, то активація W реактивує Q і R [25], подібно до асоціативної пам'яті. Це призводить до часового групування подій і може лежати в основі формування епізодичної та автобіографічної пам'яті.
- Перебір варіантів послідовностей. Коли вивчені послідовності АБВГ та АБЯЮ, то у точці розгалуження АБ спостерігається підвищена активність. Якщо вхід до нейронної мережі припиняється, то вона робить вибір між ВГ та ЯЮ.

Багато було створено моделей, що намагалися відтворити ці біологічні ефекти, проте, поки не існує такої, щоб пояснювала все [15]. Проблема, що ми не розуміємо всі обчислювальні та біологічні деталі. До таких попередніх моделей відносяться:

- Синактивні ланцюги (Synfire chains), в яких два нейрони з'єднуються, якщо вони активуються друг за другом. Тоді активація першого нейрона в послідовності призводить до активації другого і т.д. $\bigcirc \Rightarrow \bigcirc \Rightarrow \bigcirc \Rightarrow \bigcirc \Rightarrow$.
- Полісинхронізація [78] - використовує ідею часової затримки передачі спайку, що відрізняється для різних аксонів. Таким чином, ті нейрони, на які прийде активація одночасно, будуть активні і закодують наступні елементи послідовності. Через велику різноманітність довжин аксонів, а отже

затримки сигналу, завжди знайдеться унікальна комбінація нейронів що закодує елемент послідовності.

- Спайкові нейронні мережі з правилом навчання STDP показують здатність до запису та відтворення послідовної активності нейронів [3]. Проте, цей підхід має ті ж недоліки що й два попередні: нові послідовності перезаписують старі, і тому такі нейронні мережі здатні записувати дуже обмежену кількість послідовностей.
- Рекурентні нейронні мережі, що навчаються через зворотне поширення похибки в часі. Цей підхід показує найкращі практичні результати в роботі з послідовностями, зокрема розпізнавання голосу, чи переклад тексту. Це стало можливим завдяки переходу до більш складних моделей, зокрема LSTM або GRU , що можуть враховувати довготермінові залежності. Проте, такий підхід навчання не є біологічно правдоподібним.
- Ієрархічна часова пам'ять (HTM) [62, 30, 31], яка найближче по духу до даної роботи. Вона також використовує модель нейрона, що враховує дендрити в якості детектора збігів, та кодує інформацію у розріджено розподілених представленнях. Це дозволяє записувати велику кількість послідовностей. Проте, цей підхід використовує організацію нейронної мережі з міні колонками, де номери активних нейронів в колонці унікально кодують контекст минулої активності послідовності. Через це виникають проблеми з часовою інтеграцією послідовності (temporal poller) та з чутливістю до шумів.

В даній роботі детальніше досліджено прогнозування послідовностей та часову інтеграцію. Можливий механізм прогнозування в біологічних нейронних мережах виникає завдяки нелінійності дендритів [15, 20]. Красивий експеримент [21] показав, що упорядкована локальна активація з кінця дендрита до тіла нейрона викликає більшу активацію, ніж у протилежному напрямку. Та-

кий результат був теоретично передбачений з моделі Ралла, але це вимагало нереалістично довгого дендрита. Проте, тепер показано, що селективність нейрона до різних напрямків (від центру чи до центру) досягається через нелінійну активацію рецепторів NMDA та збільшення імпедансу (аналог опору мембрани клітини для змінних струмів) з більш високою відстанню від центру. Отже, послідовна активація дендриту з кінця до центру деполяризує дендрит більше, та NMDA рецептори активуються сильніше. Цей експеримент показує, що нейрон може виконувати більш складніші обчислення, ніж це вважалося раніше, а саме розпізнавати просторово-часові послідовності.

Крім того, згідно з теоретичними розрахунками [14], дендрити можуть виявляти та диференціювати послідовності на масштабі часу ~ 1 секунда, що важливо для більшості поведінкових задач. Такий порівняно великий масштаб виникає за рахунок поширення хімічних хвиль, наприклад Ca^{2+} . І тому подальша активація вздовж дендрита може підтримувати хвилю і спричиняти збільшення активності самої клітини (Рис. 3.1а). Це добре узгоджується з нещодавнім відкриттям великого масштабу часу допустимих слідів (eligibility traces), знайдених в корі мозку [65] та гіпокампі [17]. Це свідчить про тривале збереження активації дендритів, що служить основою для подальшої інтеграції послідовності.

Інше свідчення, що нейрон розпізнає просторово-часову активацію, випливає з недавнього експерименту, в якому було виміряно рецептивні поля нейронів у звуковій корі гризунів [154]. На рис.3.1б наведено приклад одного з таких полів. Червоні крапки представляють збуджувальні синапси, а сині – гальмівні. Найважливішою інформацією є те, що рецептивне поле розподілене в часі, дуже локалізоване на певних частотах та розріджене, що означає, що лише невелика частина частот визначає активацію нейрона. Головний результат показаний авторами роботи, що подібні рецептивні поля також! формуються в штучній нейронній мережі, оптимізованій для передбачення наступних елементів.

Така властивість розпізнавати просторово-часові патерни одним нейроном

покладена в основу моделі прогнозування послідовності представлений далі. По суті, використовується вже давно відомий підхід в time-delayed мережах, коли послідовність векторів перетворюється в один довгий вектор і подається на вхід мережі. Тільки в даній роботі використовується розріджено розподілене представлення даних та моделі, що підкріплені біологічними даними.

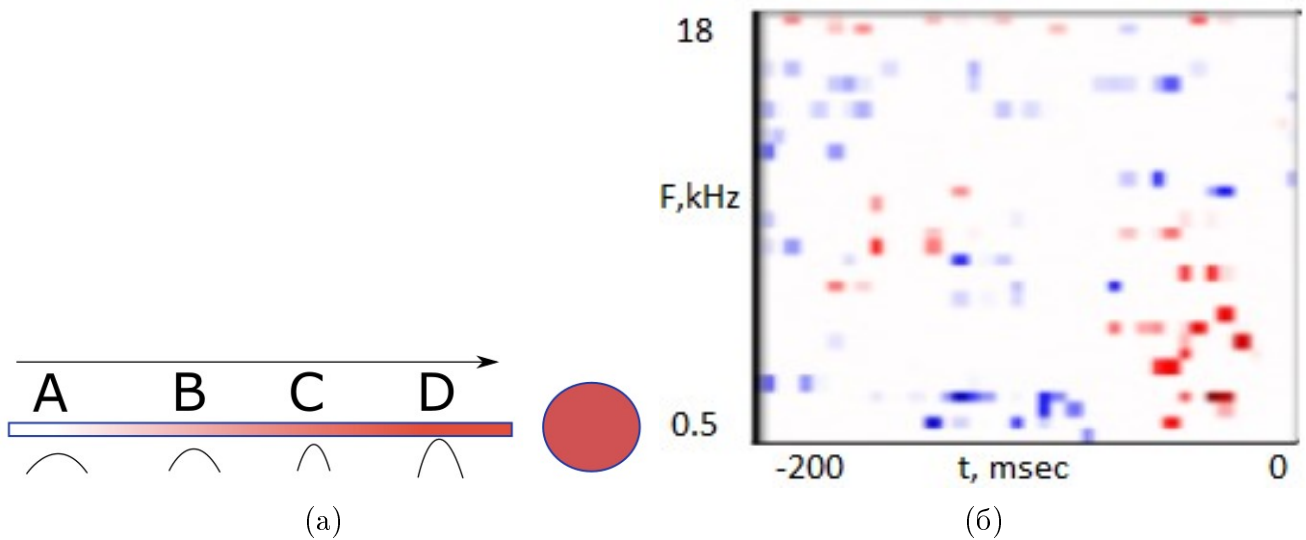


Рис. 3.1: а) Поширення хімічних хвиль вздовж дендриту б) Просторово-часові рецептивні поля гризунів в звуковій корі мозку A1 [154]. Показує, що індивідуальний нейрон дуже селективний для певних частот і часу коли вони активні.

Часова інтеграція

Така здатність розпізнавати просторова-часову активність мережі може бути використана для прогнозування: нейрон активується, що до того як прийшов відповідний сенсорний стимул. Також, є ще інший розповсюджений ефект, коли певна популяція нейронів активна впродовж всієї послідовності. Тобто, на одному нейронному шарі розвиваються послідовні просторові патерни, що відтворюють сенсорний вхід, а на іншому шарі активується популяція, що кодує всю послідовність. Тобто, відбувається представлення всієї послідовності в іншій нейронній популяції. Така часова інтеграція значно гірше досліджена, адже потребує спостереження за активністю одразу багатьох нейронів на довгому часовому інтервалі. Проте є певні припущення, яким чином нейрони зв'язують елементи послідовності в одне представлення. Одне з таких є алокація пам'яті в

нейрони зі збільшеною здатністю до активації (збудливістю, від англ. excitability) для зв'язування активації в різні моменти часу. Нещодавно було показано, що активні нейрони зі збільшеною здатністю до активації, тобто готовністю генерувати потенціал дії, з більшою ймовірністю будуть активуватися знову і тому закодують інформацію про наступний сенсорний вхід [147, 151]. Такий ефект спостерігається на часовому масштабі кількох годин-днів і припускається, що він відповідає за формування епізодичної пам'яті. Тобто, сусідні в часі події представляються схожими нейронними популяціями, а отже при нагадуванні однієї події, активуються нейрони, що представляються наступні події. Термін алокація (allocate) означає вибір нейронів, які будуть активуватися (0 або 1) на вхідний стимул. Здатність до активації означає близькість напруги на мембрані нейрона до порогового значення, що збільшує ймовірність генерації потенціалу дії (0 або 1). Таким чином, більша здатність, збільшує ймовірність активації, а отже визначає алокацію пам'яті (представлення наступних елементів послідовності).

Також, нещодавні гарні експерименти показали, що штучно збільшуючи або зменшуючи здатність нейрона до активації, через оптогенетичні методи, чи через збільшення концентрації молекул факторів росту BDNF, чи контролюючи концентрацію іонів, можна визначити чи буде нейрон брати участь в кодуванні інформації [81]. Експериментальні результати показують, що при активації нейрони також збільшують свою здатність до активації, а отже, вони, швидше за все, знову будуть активними в майбутні моменти часу і візьмуть участі в представленні інформації. На даний момент, це одна з найкращих гіпотез, як формуються епізодичні спогади з життя.

В даній роботі припускається ідея, що такий механізм алокації відповідальний за інтеграцію послідовності і на часовому масштабі порядку секунд. Цей масштаб часу особливий, так як на ньому є багато важливих для поведінки стимулів, наприклад звук, слова, моторика, візуальний рух. Для перевірки цієї ідеї запропоновано модель часової інтеграції далі в цьому розділі.

3.2 Навчання послідовностей. Постановка задачі та основні підходи

Властивості просторова-часових патернів можна формалізувати в навчання послідовностей, яке розділяється на три складові [163]: прогнозування, інтеграція, генерація.

Задача прогнозування ставиться як пошук відображення $f : \{x_t\} \rightarrow x_{t+1}$, яке на вхід отримує послідовність елементів $\{x_t\} = x_1 x_2 \dots x_t$, а на виході видає прогноз наступного елементу x_{t+1} .

Подібна задача – інтеграція послідовності як пошук відображення $f : \{x_t\} \rightarrow y$, яке на вхід отримує послідовність елементів $\{x_t\} = x_1 x_2 \dots x_t$, а на виході видає представлення всієї послідовності y (подібно до часової інтеграції як писалося вище).

Задача генерації послідовності зворотна до інтеграції і ставиться як пошук відображення $f : y \rightarrow \{x_t\}$. В біологічних організмах це використовується для генерації моторики (як послідовності рухів $\{x_t\}$) від команди (як активної популяції нейронів y).

Ці задачі насправді є нероздільними і реалізуються в ієрархічно організованих нейронній мережі. У спрощеному вигляді це можна пояснити на рис. 3.2, де представлено два рівня x та y . Нехай вхідний сигнал надходить на шар x , де x_i представлення елементу послідовності в час t_i . Шар y інтегрує активність з шару x так, що y_i є представленням на довшому часовому проміжку. При сприйнятті вхідної послідовності, наприклад ДоРеМіФа, активуються різні нейрони, які утворюють між собою зв'язки ($x \rightarrow x$, $x \rightarrow y$, $y \rightarrow y$, $y \rightarrow x$). При повторенні вхідних звуків ДоРе утворені зв'язки $x \rightarrow x$ та $y \rightarrow x$ активують нейрони, що кодували МіФа, тим самим передбачаючи наступну активність. Зв'язки $x \rightarrow y$ інтегрують активність в часі. Наприклад, x_1 кодує До, x_2 кодує Ре, а y_1 кодує ДоРе. Це інтуїтивне пояснення значно спрощене і не включає такі де-

талі як: запам'ятовування контексту (якщо були послідовності СіДоРеМі та ФіДоРеЛя, то передбачення на основі двох останніх нот не розрізнить Мі та Фа), представлення тривалості стимулу та інтервалів (тривалість ноти До та час між нотами), стиснення даних (представлення в умовах коли було багато вхідних послідовностей, але немає ресурсів всіх їх записати, і тоді потрібно вибирати ті, які найбільш важливіші). Цей простий приклад з нотами базуються на результатах багаторічних досліджень біологічних нейронних мереж, в який справді інформація обробляється ієрархічно на різних часових та просторових масштабах [60, 169]. Наприклад, рецептивні поля первинної зони звукової кори А1 можуть мати селективність в діапазоні 1с (шар x), в той час як вищі рівні, наприклад А4, можуть бути селективні на масштабі 5с (шар y).

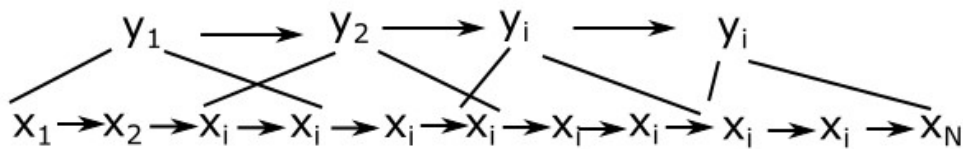


Рис. 3.2: Простий приклад ієрархічної обробки послідовностей

Основні підходи

На сьогодні виділяють два основні підходи в навчанні послідовностей: через розподіл ймовірностей та завдяки нейронним мережам. Моделі порівнюється за такими критеріями як точність прогнозування; кількість пам'яті, що займає модель; та швидкість прогнозування. Порівняння відбувається за допомогою стандартних наборів даних, наприклад слова з 100мб Вікіпедії.

Ймовірнісні моделі намагаються побудувати спільний розподіл ймовірностей елементів послідовності. Розглянемо наступну задачу. Нехай є послідовність $\{x_t, t = 1 : T\}$, де $x_i \in A$ та A -алфавіт розміру $|A| = N$. Спільний розподіл ймовірностей (СР) довжиною n визначається як $p(x_{i+s}, x_{i+s-1} \dots p_i), \forall i = 1 : T - n$, де $n \ll T$. Наприклад, якщо $\{x_t\}$ послідовність букв, то для $n = 2$ СР представляє бі-грами, ймовірність зустріч двох суміжних символів. В загальному

випадку маємо n -грами. Маючи СР робити передбачення можна доволі легко:

$$p(x_{i+1}|x_i, \dots x_i - k) = \frac{p(x_{i+1}, x_i, \dots x_i - k)}{p(x_i, \dots x_i - k)}$$

, де чисельник та знаменник отримується з СР шляхом маргіналізації (підсумуванні по всіх інших елементах). Головна проблема СР в експоненційному збільшенні розміру матриці зі збільшенням n , яка має розмір $\underbrace{N \times N \times \dots \times N}_n$ і потребує N^n чисел для опису. Для $N = 30, n = 10, N^n \approx 10^{15}$ - це величезне число. Проте матриця сильно розріджена, більшість чисел нульові. Розрідженість виникає через недостатню кількість даних, щоб все заповнити, або дані структуровані, коли деякі події ніколи не зустрічаються разом. Для зменшення розміру матриці застосовують різні підходи, рахують розподіл не суміжних символів, а з певним кроком, що називають скіп-грами. Або створюють ієрархічне представлення даних, коли підпоследовності $\{x_r\}$ присвоюють новий елемент y_i (який вже кодує, припустимо, слово), і рахують розподіл ймовірностей на новому рівні меншої розмірності. Так само можна продовжити і на рівень речень, або всього тексту, де нові змінні кодують контекст. Проте, з таким підходом виникають проблеми, а саме, з вибором підпоследовності якій присвоїти новий елемент.

Варто звернути увагу на найбільш популярний ймовірнісний підхід – приховані марківські моделі. Марківський процес порядку k робить припущення про залежність передбачення наступного елемента тільки від k минулих подій $p(x_{i+1}|x_i, \dots x_i - n) = p(x_{i+1}|x_i, \dots x_i - k), n > k$. У звичайному випадку спостерігачу надходять елементи последовності x_i в кожен момент часу і він робить припущення про джерело, що їх породжує, тобто будує його модель. Інакше кажучи, будує спільний розподіл ймовірностей джерела. Приховані марківські процеси роблять припущення, що дані генерують кілька джерел, і ймовірність, що дані надходять з певного джерела, визначається марківським процесом. Тобто вибір звідки йдуть дані залежить тільки від того звідки йшли дані за ми-

нулий крок (у найпоширенішому випадку марківського процесу порядку 2). І тоді спостерігач може будувати не спільну усереднену модель, а модель кожного джерела, що буде значно точніше та даватиме кращі передбачення. Під таке формулювання підпадають багато практичних задач, як розпізнавання звуку, рукописного тексту або послідовності нуклеотидів. І довгий час приховані марківські моделі показували найкращі результати. Проте вони мають свої недоліки. Якщо кількість прихованих станів S (джерел), то обчислювальна складність складає $o(S^2)$ згідно алгоритму Вертербі. Більше того, ймовірність переходу між станами залежить тільки від минулого стану. Хоча й можна збільшити залежність на більш дальні події через створення нового простору прихованих станів, що комбінує минулі стани (декартовий добуток минулих станів), проте розмірність експоненційно виросте і від цього зросте обчислювальна складність [176] .

Ще цікавим способом працювати з СР великої розмірності - робити його компресію в новий простір \hat{p} значно меншої розмірності. Так як оригінальний СР сильно розріджений, а отже має низьку ентропію, і тому може добре стискуватися. І далі робити процедуру оновлення вже в новому просторі. Щось подібне виникає в прихованих марківських моделях, а також моделях нейронних мереж, як штучних так і біологічних [10].

Станом на сьогодні рекурентні нейронні мережі мають найкращу ефективність в навчанні послідовностей. Вони показують найкращі результати в багатьох областях, таких як розпізнавання голосу та рукописного тексту, перекладу тексту, генерації опису зображення чи відео. Вони працюють краще, тому що можуть відслідковувати довготермінові залежності на відмінно від прихованих марківських моделей. Це особливо справедливо для нейронних мереж LSTM/GRU, які мають спеціальну архітектуру саме для збереження залежностей між подіями розділених тривалим інтервалом. Такий успіх зумовив їх широке комерційне застосування, що раніше бувало доволі рідко з нейронними мережами. Проте і ці моделі мають ряд недоліків. Основний – це необхідність

великої кількості розмічених даних для навчання з учителем, а також великі обчислювальні ресурси. Також, вони поділяють недоліки в загальному від методів навчання як оптимізації певної функції, що немає інтуїтивного розуміння роботи алгоритму, тобто дуже складно розв'язок складної задачі розкласти на розв'язок менш складних задач. Все працює як чорний ящик, параметри якого визначає мінімізація похибки.

Задача передбачення тісно пов'язана із задачею стиснення. Дві ключові складові стиснення даних – це побудова моделі та вибір способу кодування. Теорія кодування вже досить добре розроблена і є хороші алгоритми як найкраще кодувати дані, щоб передавати з мінімум втрат (наприклад коди Хоффмана, арифметичні коди чи коди з малою щільністю на перевірку парності. Проте, побудова моделі - складна невирішена задача, яка можливо і не має загального рішення. Під моделюванням частіше всього розумію побудову розподілу ймовірностей даних. І як уже було сказано раніше, через велику розмірність практично це зробити не вдається, тому застосовують різні евристики під конкретні задачі. "На даному етапі стиснення даних це одночасно і мистецтво і проблема штучного інтелекту" як написано в [108]. Якщо побудована система буде добре стискати дані, значить вона і буде робити хороші передбачення.

Також задача навчання послідовностей виникає в навчанні з підкріпленням, коли агенту потрібно вивчити послідовність дій, що максимізує винагороду [173]. Класичний підхід до рішення задачі є алгоритм TD(temporal difference) навчання (навчання часових різниць).

Існує ще інший підхід - пошук правил або інваріантів. В такому випадку, шукаються правила, що генерують послідовність, які можуть застосовуватися для зовсім нових випадків. Наприклад, можна навчати систему передбачати рух кинутого під кутом м'яча. Назбирати дані для експериментів від різними кутами, вагою м'яча, швидкості вітру, і т.д., і побудувати модель з багатьма параметрами, яка буде передбачати траєкторію для всіх інших параметрів. Або можна записати рівняння руху (що робиться у фізиці) в компактній аналітичній

формі для всіх можливих параметрів. Таке рівняння має найкоротший опис (див. принцип довжини мінімального опису (minimum description length), що подібне до принципу Окамі). В той час нейронні мережі та інші статистичні моделі потребують багато параметрів, що лише наближають до справжнього аналітичного опису.

Як видно, навчання послідовностей досить актуальна задача з багатьма запропонованими алгоритмами. Тим не менш, в загальному виділяють наступні проблеми в навчанні послідовностей [163]:

1) Навчання довгочасових залежностей. Часто на майбутнє впливають події в далекому минулому. Щоб правильно робити передбачення потрібно багато ресурсів і кращі алгоритми, що відслідковують всю історію послідовності.

2) Ієрархічне структурування послідовності. Більшість реальних послідовностей мають композиційну структуру, тобто послідовність складається з підпослідовностей, які теж можуть складатися з підпослідовностей. Відкритим питанням є як виділяти ці підпослідовності. Вирішення цієї проблеми допоможе стиснути послідовність, а також допоможе "скоротити" часові залежності.

3) Проблема хаотичних послідовностей, які часто зустрічаються в реальному середовищі, наприклад передбачення часових рядів.

Варто зазначити, що біологічні нейронні мережі в певній мірі вирішують дані проблеми, і відкритим питанням залишається яким чином це робить розподілена система (нейронна мережа) через самоорганізацію.

3.3 Модель прогнозування послідовностей на основі сігма-пай нейрона та розрідженого представлення даних

В залежності від типу даних задача прогнозування може розділятися на дві:

1) прогнозування часових рядів у випадку, якщо послідовність x_t описується

числами; 2) прогнозування елемента послідовності, коли послідовність x_t - сим-
воли з обмеженого словника.

В даному розділі буде розглядатися друга задача.

Формалізація моделі

В даній роботі представляється біологічно подібна модель прогнозування по-
слідовності. Використовується ідея, що індивідуальний дендритний сегмент мо-
же працювати не тільки як детектор просторових збігів (комбінація одночасно
активних нейронів), але й часових (комбінація не одночасно активних нейронів).
Вище вже писалося, що в біологічних експериментах показано здатність дете-
ктувати просторово-часові патерни одним нейроном. Далі пропонується модель,
де кожен нейрон працює як детектор просторово-часових збігів. Елементи по-
слідовності представлені як розріджені бінарні вектори в дискретних моментах
часу. Розріджені, нагадаємо, значить, що кількість активних нейронів значно
менше всієї кількості $s = a$

$N \ll 1$. Задача прогнозування зводиться до моделі просторової пам'яті з дру-
гого розділу. Для цього послідовність бінарних векторів за останні T часових
кроків конкатенується в один великий вектор. І створюються асоціації між бі-
нарним вектором розміром $T \cdot N$ з вектором розміром N , що представляє на-
ступний елемент послідовності. Важливо, щоб участь в асоціації взяли активні
нейрони з попереднього кроку (які є частиною вектору розміром $T \cdot N$), іна-
кше вихідний вектор буде активуватися патерном, що передбачає не наступний
елемент, а більш майбутній. Цей підхід схожий на нейронні мережі з часом за-
тримки, де контекст для передбачення встановлюється нещодавньою історією
активності послідовності, але відмінність полягає у використанні розріджено
розподілених представлень з більш біологічно подібною моделлю нейрона.

Написане вище можна формалізувати в наступну модель. Спершу, вхідний
сигнал $z(t_j) \equiv z_j$, вектор розміром M перетворюється на розріджений вектор

$x(t_j) \equiv x_j$ розміром N високої розмірності. Для цього використовується алгоритм випадкової проекції $x_j = kWT A(\sum_j w_j z_j)$, де $w_j \in \{0, 1\}^{N \times M}$ - випадкові бінарні ваги і функція $kWT A(\cdot)$ повертає $a \ll N$ найбільш активних нейронів. Результат виглядає, наприклад так:

$x_j = \{000000100000000001000000000000000000000000100\}$. Нещодавно було показано, що подібна процедура кодування використовується в нюховій області мух і обчислювально відповідає алгоритму локально-чутливому хешуванню (local sensitive hashing)[33].

Далі ініціалізується нульовий вектор v_j розміром $T \cdot N$ і в останні елементи записується значення x_j . Обраховується нове значення x_{j+1} , яке асоціюється з v_j за алгоритмом в розділі 2. Далі оновлюється вектор v_{j+1} шляхом здвигу вектору x_j та записування на останні елементи вектор x_{j+1} . Далі, оновлений вектор $v_{j+1} = [00...00x_jx_{j+1}]$, асоціюється вже з x_{j+2} . Розглядається два способи асоціації. В першому, з кожним елементом x_j утворюється k_1 зв'язків, тому сумарно один нейрон утворює $k_1 \cdot T$ зв'язків, В другому, з останнім елементом x_{j+T-1} утворюється m зв'язків, а зі всім іншими випадково розподілено k_2 зв'язків, тому сумарно один нейрон утворює $m + k_2$ зв'язків. Процедура повторюється для всієї послідовності. Схематично утворення зв'язків одним нейроном показано на рисунку.3.3

Представлена модель відрізняється від попередніх тим, що вхідний сигнал перетворюється в розріджений вектор високої розмірності та використовується функція активації нейрона сігма-пай. Подібні ідеї були представлені раніше в [19]; однак, тут використовуються інші функції активації та правила навчання, що дозволило отримати набагато більшу ємність пам'яті. Також варто зазначити, що дана модель є більш біологічно подібно, так як використовується локальне правило навчання та конкатенація векторів в один, що відповідає часовій інтеграції одним нейроном.

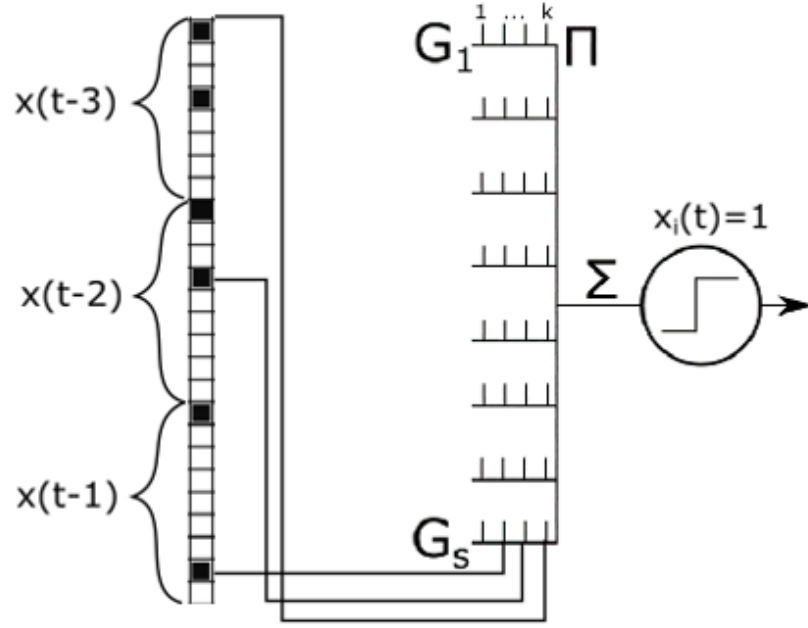


Рис. 3.3: Ілюстрація ідеї з'єднання популяцій в часі

Теоретичний аналіз та результати симуляції

Виведення ємності пам'яті переходів

Теоретичний аналіз даної моделі подібний до аналізу просторової асоціативної пам'яті, представленого в другому розділі, з різницею у врахуванні часового вікна. Максимальна кількість пам'яті переходів має такий самий вираз як і 2.4, тільки $s' = \frac{a'}{N'} = \left(\frac{C_a^k}{C_N^k}\right)^T$.

Нехай мережа розміром N , де кількість активних нейронів a , розрідженість $s = \frac{a}{N} \ll 1$, кількість синапсів в одному кластері k , кількість часових кроків T , що впливають на активацію нейрона. Ця задача є еквівалентною до формування асоціацій в нейронній мережі розміром $N' = (C_N^k)^T$ і активацією $N' = (C_a^k)^T$, де C_a^k біноміальний коефіцієнт. Позначимо w як матрицю віртуальних зв'язків розміром $[N, N']$, яка навчається як $w^i = x^i f(\vec{x})$ де $f(\vec{x})$ повертає один випадково вибраний елемент посеред a' в мережі N' . Насиченість зв'язків для нейрона визначається як відношення кількості ненульових зв'язків до загальної кількості $s_w^i = \frac{|w^i|}{N'}$. Кожен крок навчання збільшує насиченість ваг на $s_0 = \frac{1}{N'}$. Після

В записаних події, насиченість:

$$s_w = 1 - (1 - s_0)^{Rs} \quad (3.1)$$

Ймовірність помилкової активації нейрона визначається як:

$$p(\tilde{x}_m^i = 1 | x_m^i = 0) \equiv p = 1 - (1 - s_w)^{a'}(1 - s) \quad (3.2)$$

де \tilde{x}_m^i передбачення для нейрона i в момент часу m . Встановлюємо параметр точності $\varepsilon = 0.01$, що визначає максимальний відсоток помилкових передбачень тобто менше ніж 1% клітин можуть бути помилково активовані. З урахуванням заданої точності можна порахувати максимальну кількість переходів R_{max} , що нейронна мережа може записати і точно передбачити. Також беручи граничний випадок $s_0 \rightarrow 0$ ми отримуємо:

$$R_{max} = \frac{\varepsilon}{s(1 - s)s'} \quad (3.3)$$

, де $s' = \frac{a'}{N}$. З цього теоретичного аналізу маємо важливий висновок, що для розрідженої активної мережі виявлення точного збігу в часі працює як комбінаторне розширення розмірності мережі, в якій легше розділити різні патерни. Максимальна кількість пам'яті переходів послідовності залежить від величини розрідженості мережі та кількості синапсів у кластері. Збільшення розмірності призводить до збільшення розрідженості, що призводить до збільшення R_{max} . В граничному випадку коли одна клітина активна, розрідженість найбільша, але в цьому випадку використовується локальне кодування, тому мережа може записати лише N переходів. Нижня межа на розрідженість визначається через нерівність $C_N^a > R_{max}$.

Визначення ємності для різних параметрів

Далі досліджено наскільки ємність моделі залежить від параметрів. На рис. 3.4а. показано, що чим довший час історії T , який впливає на нейрон, тим вища

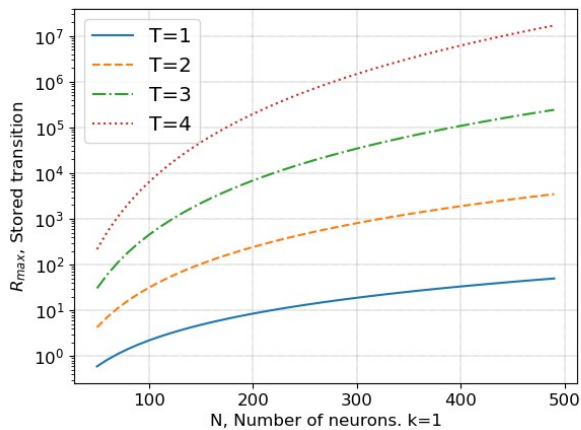
ємність. Крім того, чим більше нейронів з одного моменту часу з'єднуються у кластер, тим більша ємність. Такі результати впливають оскільки при випадковій активації збіг активності трьох нейронів набагато рідкісніший, ніж збіг двох. Тому, активація навіть кількох нейронів служить хорошим представленням історії активності. Цікаво, що для певної комбінації параметрів нейронна мережа може передбачити 10^{10} переходів. Таке величезне число вимагає величезної кількості синапсів і практично та біологічно не є реалістичним.

Модифікація моделі

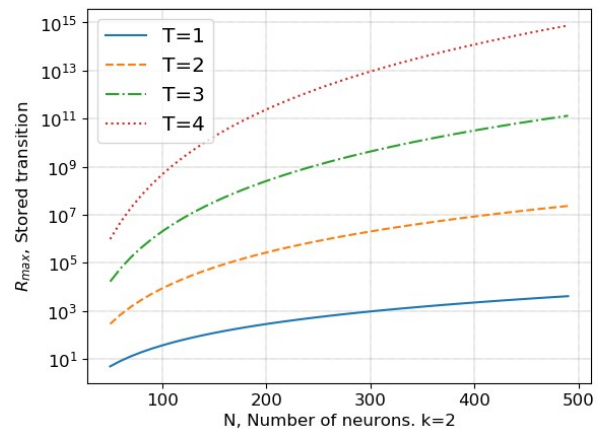
Як писалося вище розглядається два способи активації. Результати для першого способу, коли з кожною подією створюється фіксована кількість зв'язків вже наведені. Мотивація другого способу полягає в тому, що попередній спосіб створює біологічно нереалістично багато зв'язків на нейрон. Тому запроваджується обмеження кількості зв'язків до $m + k_2$ (з передостаннім та + всіма іншими елементами послідовності), що зменшує ємність переходів. Але оскільки зв'язки розподілені у просторі та часі система все ще має хорошу ємність. У випадку зменшеної кількості зв'язків збільшення розмірності відбувається до $N' = C_N^{k_2} C_{N(T-1)}^m$ та активації до $a' = C_a^m C_{a(T-1)}^{k_2}$ де m кількість зв'язків до події на передостанньому кроці, та k_2 кількість зв'язків до подій в проміжку T-2 кроків назад.

На рис.?? представлені результати ємності для різних розмірів мережі для різних часових вікон з $m = 1, k_2 = 2$, та $a = 20$ - середня кількість активних клітин. На правій осі представлена кількість синапсів на нейрон, що відповідають пунктирним кривим. Ємність не залежить від часового вікна T і збільшується з розміром мережі. Для $N = 1000$ існує 4000 синапсів на нейрон, що є біологічно правдоподібним, а кількість переходів дорівнює 10^5 . Якщо взяти послідовності як слова з середньою кількістю символів 7, мережа може записати близько 14000 можливих слів.

Тим не менше, ці модифікації знаходяться на ранній стадії досліджень. Результати на реальних наборах даних, таких як TIMIT або Penn tree bank (PNB),



(а)



(б)

Рис. 3.4: Їмність пам'яті переходів послідовностей для різних параметрів. (а) Їмність з $k = 1$, так що один синапс створюється за подію. Їмність для великих мереж досягає мільйонів переходів. б) їмність з $k = 2$. У цьому випадку нейрон краще розпізнає події, тим самим він може запам'ятати набагато більше переходів.

повинні бути отримані в подальших роботах. Однак головна мета моделі виконана - запропоновано можливий спосіб прогнозування послідовностей в біологічних нейронних мережах з урахуванням сучасного розуміння дендритних обчислень.

Важливість розріджено розподілених представлень

Переваги представленої моделі виникають з розріджено розподіленого представлення вхідних даних [82, 83]. Розрідженість нейронної активації дає високу місткість пам'яті. Теоретична формула максимальної кількості пам'яті переходів містить розрідженість в знаменнику, тому вища розрідженість призводить до вищої пам'яті. Як зазначав Барлоу [10], біологічна нейронна мережа повинна кодувати ймовірності подій, щоб мати можливість застосовувати подібну обробку як баєсівський підхід [88, 28].

Розрідженість має вирішальне значення для представлення ймовірностей, оскільки щільне представлення має багато активних нейронів, і патерни активації сильно перекриваються і важко розрізняються. Наприклад, якщо дві послідовності однакові на інтервалі T , то модель не може правильно передбачити

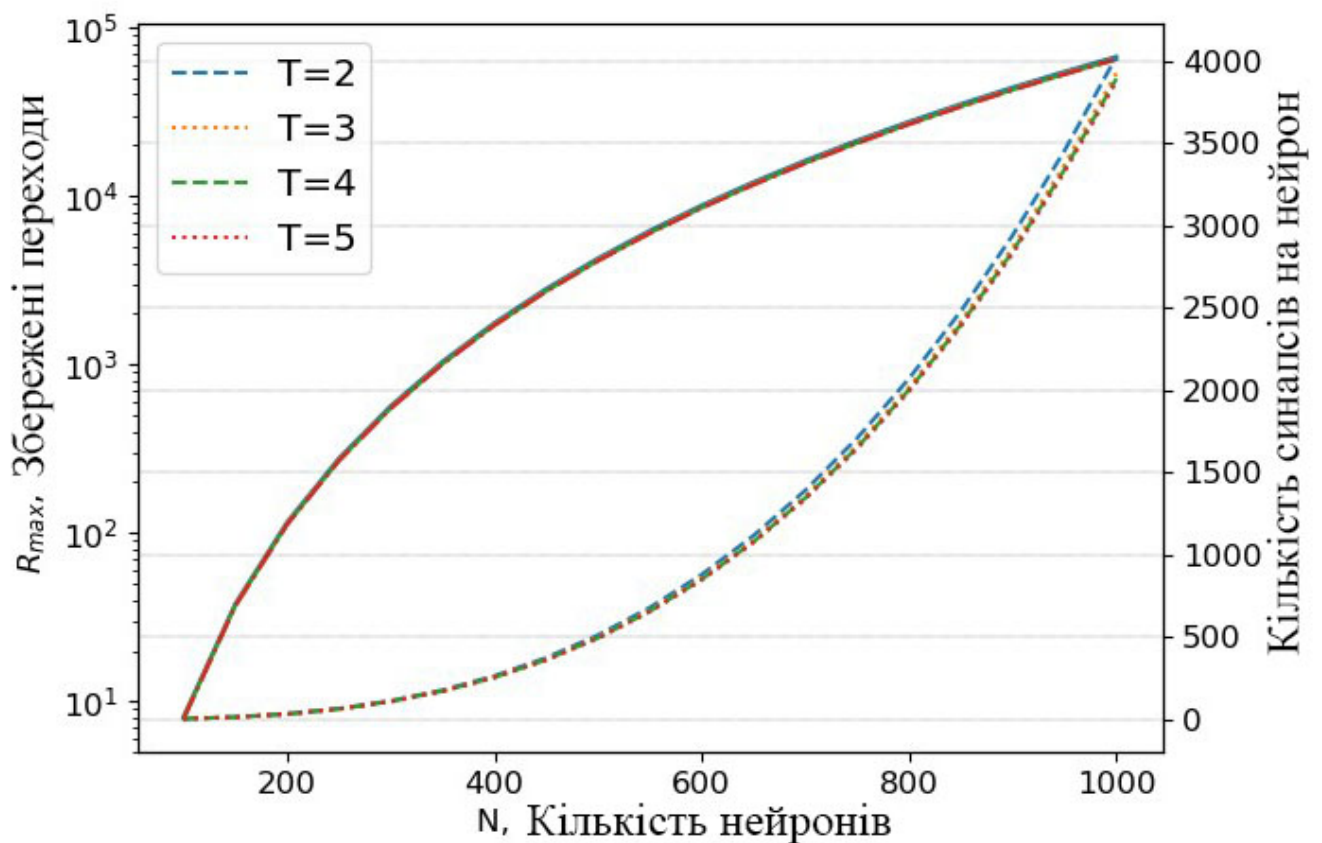


Рис. 3.5: Максимальна кількість переходів послідовностей для різних розмірів мережі та кроків часу показана як верхня крива. Нижня крива показує кількість зв'язків на один нейрон

наступний елемент. Для послідовностей ABCD та ABCR для $T=3$, мережа сплутає D та R. Але якщо перша послідовність подавалася раніше тричі, а друга всього лише, чи не має система передбачити D, як більш ймовірний наступний елемент? Гіпотетично, якщо перша послідовність була частіше, тоді і зв'язків має утворитися більше, і тому популяція D має бути більш активною, що виграє в популяції, що кодує R. Для цього потрібно внести незначні зміни до моделі, щоб мережа почала відтворювати ймовірності подій, проте залишається проблема з правилами видалення зв'язків, або коли кількість можливих зв'язків до запису обмежена. Тоді потрібно вибирати, які зв'язки можна видалити, а які залишити.

Також інше значення розрідженості полягає в здатності до узагальнення,

тобто прогнозувати не тільки для раніше відомих послідовностей, але і для нових. Використання випадкової проекції має властивість збереження схожості для вхідних даних. Два подібні вхідні сигнали кодуються подібними представленнями. Якщо в двох патернах є спільні нейрони, тоді вони активуються частіше, і тому створюють більше зв'язків і при передбаченні активуються в першу чергу. Таким чином розмиваються особливості кожного патерна і вибираються загальні подібні риси.

Метод навчання запропонований в моделі припускає, що в біологічних мережах структурна пластичність відіграє ключову роль для вибору місця розміщення синапсів у конкретних дендритних місцях. Важливо, що зв'язок характеризуються не тільки силою, як загально прийнято, але й позицією на дендритному дереві. Це допомагає розпізнавати правильні комбінації просторово-часової активності. Дана модель черпає натхнення з подібних моделей запропонованих раніше командою Numenta [6, 62]. Вони використовують нейрон з багатьма дендритами, які діють як детектор збігів і кодують події у розріджено розподілене представлення. Основна відмінність полягає в тому, що у представленій моделі не потрібна спеціальна структура мережі з міні колонками, контекст послідовності записується через часове вікно. Це дозволяє отримати збереження схожості контексту для схожих послідовностей. Також, завдяки мультиплікативній функції активації отримано на порядок вищу ємність пам'яті переходів при однаковій кількості зв'язків. В моделі НТМ також використовують модель нейрона з дендритами, тільки кожний дендрит активується від зваженої суми з порівняння з порогом ($\sum w_{ijk}x_i > \theta$, активації k-го дендриту j-го нейрона). На кожному дендриті 30 синапсів, як запам'ятовують просторовий паттерн активації. В даній ж роботі показано, що не потрібно 30 синапсів, достатньо кількох, якщо використовувати мультиплікативну активацію $\prod w_{ijk}x_i$. За рахунок цього і досягається більша ємність.

Важливим недоліком запропонованої моделі є відсутність механізму видалення синапсів, що з часом призводить до їх великої кількості та погіршення

прогнозування. Основна проблема запропонованої моделі, що вона не враховує події, які відбулися поза вікном часу T . Тому потрібно використовувати в якості контексту інші величини. В наступному підрозділі пропонується модель часової інтеграції послідовності, яка може бути використана для створення нового контексту для задачі прогнозування.

3.4 Часова інтеграція послідовностей в нейрони зі збільшеною здатністю до активації

Для зручності читача, деякий матеріал, представлений далі, буде повторюватися зі вже написаним.

Огляд

Згідно з популярним підходом за Д.Марром [32], розуміння роботи нервової системи зручно представляти на трьох рівнях: обчислювальному, алгоритмічному та фізичної реалізації. Перші два відповідають на питання, які обчислення виконуються в мозку, та який їх алгоритмічний опис, що переводить вхід у вихід. Третій намагається відповісти, яким чином абстрактні принципи обробки інформації реалізуються в матерії, зокрема в біологічних нейронних мережах. Даний підрозділ намагається поєднати всі три рівні для задачі інтеграції часової послідовності, що може бути представлена як пошук відображення $f : x_t \mapsto y$ [163], де $x_i \in \{0, 1\}^N$ - елемент послідовності представлений як бінарний вектор розмірності N , $x_t = x_1 x_2 \dots x_t \in \{0, 1\}^{t \times N}$ послідовність елементів, $y \in \{0, 1\}^M$. Таке відображення $f()$ використовується в передбаченні послідовностей з довгочасовими залежностями, асоціації послідовності з послідовністю та в задачах навчання з підкріпленням. Також є численні свідчення про використання такого відображення в біологічних нейронних мережах [113, 120, 166, 121].

Далі представлено біологічно подібний алгоритм інтеграції послідовності.

Показано, що біологічний механізм алокації пам'яті в нейрони зі збільшеною здатністю до активації, у поєднанні з розрідженою бінарною активністю мережі дозволяє отримувати представлення послідовності з корисними властивостями описаними далі. Термін алокація означає вибір нейронів, які будуть активуватися (0 або 1) на вхідний стимул. Здатність до активації означає близькість напруги на мембрані нейрона до порогового значення, що збільшує ймовірність генерації потенціалу дії (0 або 1). Таким чином, більша здатність до активації, збільшує ймовірність активації, а отже представлення наступних елементів послідовності. Проведені обчислювальні експерименти із запропонованою моделлю показують збереження схожості представлення послідовностей при шумі та перестановці елементів, часову схожість та доповнення послідовності, що якісно відтворює біологічні ефекти. Запропонований механізм є біологічно подібним і може використовуватися для часової інтеграції послідовності.

Постановка задачі та модель

Математично, задача інтеграції послідовності може бути сформульована як:

$$y = f(x_0, x_1 \dots x_j)$$

Де, $x_i \in \{0, 1\}^N$ – i -тий елемент послідовності представлений як бінарний вектор розмірності N , $y \in \{0, 1\}^M$ - представлення всієї послідовності. Задача стоїть як пошук функції $f()$, що формує представлення y , яке відтворює біологічні ефекти зазначені вище, тобто збереження схожості послідовності, чутливість до порядку, доповнення послідовності. Основна мета, зрозуміти алгоритм роботи біологічних нейронних мереж через відтворення ефектів, тому розв'язок має бути біологічно подібний. Для вирішення задачі побудована нейронна мережа з трьома шарами прямого поширення (Рис 3.6). На вхід по одному елементу подається послідовність $\alpha_1 \alpha_2 \dots \alpha_T$ довжиною T . Елементи послідовності належать до скінченного словника $\alpha_i \in A, |A| < \infty$ (наприклад букви). Кожному

елементу співвідноситься бінарний код $\alpha_i \mapsto x_i$, який вперше визначається випадково з розподілу Бернуллі $x_i \sim Ber(s_x)$, з параметром s_x , що визначає розрідженість. Далі відношення між символами α_i та їх кодуванням x_i зберігається. Першим шаром мережі і слугує вектор x_i . Другий шар $z_i \in \{0, 1\}^M$ переводить вхідний вектор в бінарний розріджений вектор через:

$$\begin{aligned} z_i &= kWTA((1 + v_j) \sum_j w_{ij} x_j, k_1) \\ v_{i+1} &= v_i + \tau(z_i - v_i) \end{aligned} \quad (3.4)$$

Де $w_{ij} \in Ber(s_w)$ матриця випадкових бінарних ваг з розподілу Бернуллі розмірності $[M, N]$ при $M \gg N$, $kWTA()$ - функція (k winners take all), яка вибирає k_1 найбільших компонентів вектору і повертає їх значення одиницями, а всі інші є нульові. Ця процедура вперше запропонована в цій роботі і подібна до алгоритму випадкової проекції [141, 142], який формує розріджені бінарні вектори, де зберігається схожість між вхідними векторами x_i, x_j та їх проекціями z_i, z_j . Кодування виглядає, наприклад як $z_i = [00000101010000000000100000]$. Нещодавно було показано, що алгоритм випадкової проекції використовується в нюховій області мозку мух і обчислювально відповідає локально-чутливому хешуванню (local sensitive hashing)[33]. В даній роботі ця процедура видозмінена шляхом множення лінійної сумації в рівнянні 3.4 на збудливість нейронів v_i . Здатність до активації визначається як низькочастотний фільтр від z_i або як рухоме середнє через рівняння 3.4 з параметром τ .

Після того як кожний елемент послідовності спроектований на свій вектор z_i , визначається вектор $y \in \{0, 1\}^M$, що представляє всю послідовність:

$$y = kWTA(\sum_{i=1:T} z_i, k_2) \quad (3.5)$$

Застосовується також функція $kWTA()$ для суми векторів z_i для всіх елементів послідовності. Тільки в цьому випадку вибирається менша кількість активних нейронів, що закодують всю послідовність $k_2 \ll k_1$.

Основна ідея даної моделі в тому, що активні нейрони $z_s^i = 1$ для s -го елементу збільшують свою здатність до активації v_{s+1}^i згідно рівняння 3.4. Вища здатність нейрона призводить до вищої активації для наступних елементів z_{s+1}^i . Тобто, активні нейрони будуть більш ймовірно активні і далі, таким чином представляючи наступні елементи послідовності α_{s+1} . Ті нейрони, які були найчастіше активні увійдуть в кінцевий вектор y властивості якого і будуть досліджуватися далі.

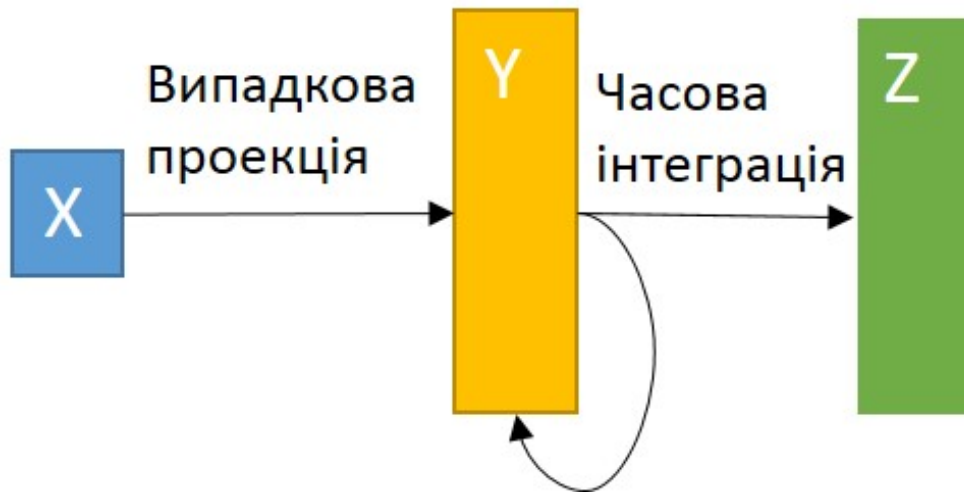


Рис. 3.6: Схематичне зображення мережі

На рис.3.7 зображено ілюстрацію ідеї моделі. Кола схематично показують активні нейрони, що кодують елементи послідовності. Нехай перший елемент послідовності А активує певну популяцію нейронів z_A , наступний елемент В активує нові нейрони $z_{B|A}$. Без введення здатності до активації, рівняння 3.4 породжує представлення елементу В незалежно від попередньої історії, але при $v \neq 0$ кожний елемент створює контекст в якому кодуються всі наступні, тому й вибрано позначення $z_{B|A}$ (В у контексті А). Перетин $z_A \cdot z_{B|A}$ закодує підпослідовність АВ (так як вектори бінарні з однаковою кількістю одиниць перетин визначається просто як скалярний добуток). Далі елемент С активує нову популяцію, яка має перетини попередніми популяціями. Нейрони в $z_A \cdot z_{B|A} \cdot z_{C|AB}$ активні найчастіше і закодують всю послідовність. У випадку зворотної послідовності СВА, елемент В буде вже кодуватися в контексті С, і тому його представлення $z_{B|}$ буде відрізнятися від представлення $z_{B|A}$. Додавання здатності

до активації в алгоритм випадкової проекції збільшує цей перетин і призводить до появи властивостей вектору y подібних до тих, що спостерігаються в біологічних нейронних мережах (показано далі).

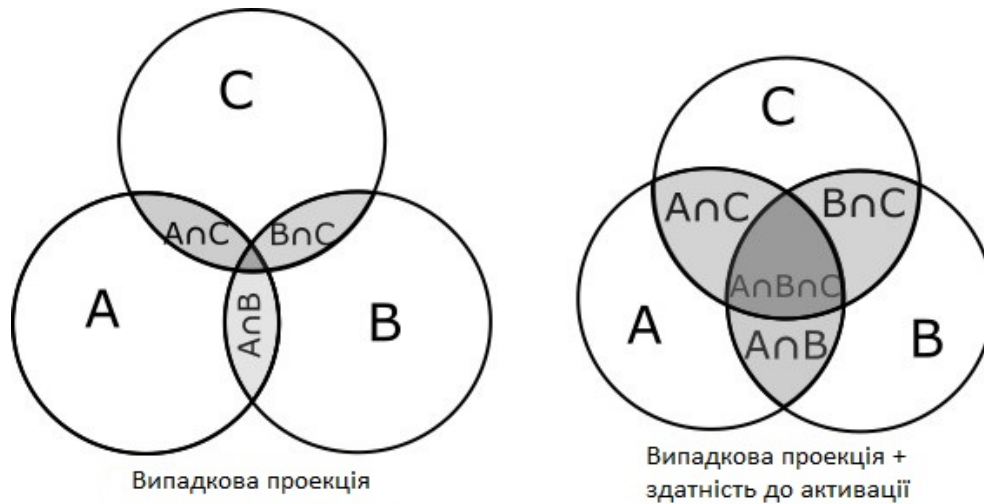


Рис. 3.7: Активні популяції від послідовності ABC та їх перетини, що кодують спільні події підпослідовності

Обчислювальні експерименти

Суть експериментів порівняти представлення (вектори y) для різних послідовностей. В якості міри схожості було взято масштабований скалярний добуток $s(y_1, y_2) = \frac{y_1 \cdot y_2}{|y_1|} \in [0, 1]$; це рівнозначно косинусній подібності так як вектори бінарні з однаковою кількістю одиниць (після $kWTA(\cdot)$).

Для всіх наступних експериментів використано наступні параметри, якщо не вказані інші: $|A| = 50$, $N = 40$, $s_x = 0.4$, $M = 1000$, $s_w = 0.2$, $k_1 = 300$, $k_2 = 50$, $\tau = 0.3$. Довжина послідовності $T = 10$. Всі експерименти проводилися за допомогою мови програмування Python.

А.Стійкість до шуму

Перший експеримент має на меті порівняти представлення y оригінальної послідовності $\{x_t\}$ та представлення \tilde{y} спотвореної $\{\tilde{x}_t\}$. Спотворення розміром n в місці i визначається як заміна n послідовних елементів послідовності починаючи з елементу i на випадкові елементи зі словника. Результати симуляції

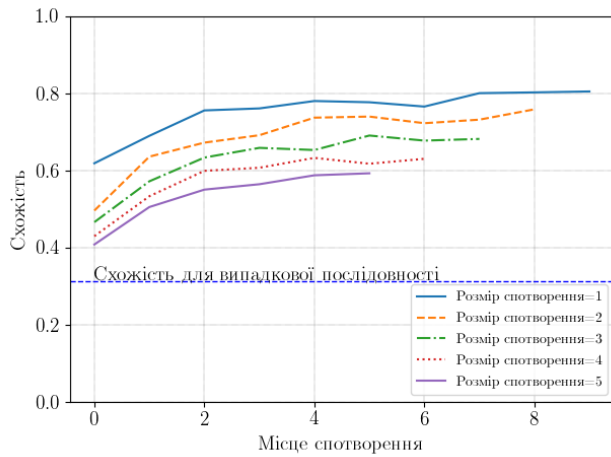
представлені на рис. 3.8а. По вертикальній осі показано схожість представлення послідовності та її викривленої версії. По горизонтальній осі відкладено місце спотворення. Різні криві показують різний розмір спотворення. Як видно, чим більше спотворення, тим менша схожість. Чим ближче спотворення до початку послідовності, тим менша схожість, що пояснюється зміною контексту для наступних елементів. Ключовим є те, що всі спотворенні послідовності мають більшу схожість, ніж довільна випадкова послідовність, яка показана штриховою горизонтальною лінією. Це свідчить про стійкість представлення послідовності до шуму.

Б. Стійкість до перестановок

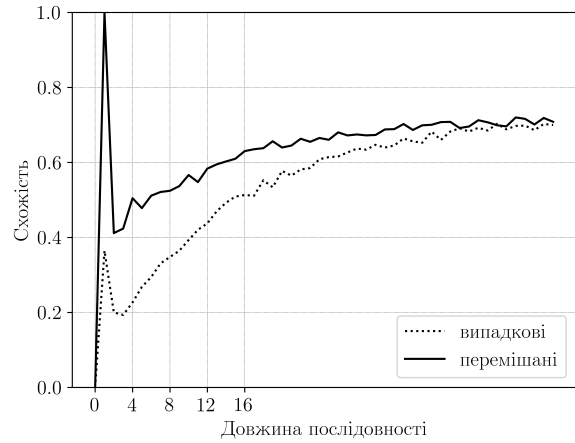
Наступний експеримент порівнює представлення послідовностей з перестановками елементів. На рис.3.8б суцільною лінією показано схожість представлення послідовності з представленням цієї ж послідовності, в якій випадковим чином перемішані елементи. Штрихом показано схожість до послідовності з випадковими елементами зі словника. Для коротких послідовностей видно, що послідовність з перемішаними елементами має більшу схожість, ніж випадкова послідовність. Тобто зберігається схожість при перестановках, але все одно, послідовності ABC та ACB представлені по різному. Розмір словника взято меншим, $|A| = 10$, щоб показати, що при збільшенні довжини послідовності, представлення випадкової та перемішаної послідовності знаходяться на однаковій відстані від оригінальної послідовності. При цьому втрачається здатність до розрізнення послідовності за даних параметрів мережі. При збільшенні розміру словника або розмірності мережі здатність до розрізнення зростає. Всі експерименти проведені з генеруванням випадкових перестановок з подальшим усередненням схожості.

В. Доповнення послідовності

Попередні два експерименти показують, що механізм алокації в нейрони зі збільшеною здатністю до активації може бути використаний для задачі доповнення послідовності. Для кращої ілюстрації цього проведений наступний експери-



(а)



(б)

Рис. 3.8: (а) Графік залежності схожості представлення послідовності з представленням спотвореної копії в залежності від місця спотворення та різних ступенів спотворення. (б) Графік залежності схожості представлення послідовності з представленням послідовності з перемішаними елементами відносно довжини послідовності. На нижній кривій зображено схожість з випадковою послідовністю.

мент: на вхід подається послідовність довжиною $T = 20$ і отримується її представлення \mathbf{y} , далі на вхід подається її підпослідовність $x_1 \dots x_t$ довжиною $t < T$ і порівнюється її представлення \mathbf{y}_t з представленням всієї послідовності \mathbf{y} . Для порівняння також знаходиться схожість з випадковою послідовністю $\tilde{x}_1 \dots \tilde{x}_t$ з елементів словника. На рис. 3.9а показано результати залежності схожості в залежності від довжини підпослідовності. Як видно чим довша підпослідовність, тим більша схожість представлення з повною послідовністю. Нижня крива показує схожість випадкової підпослідовності, яка менша порівняно з правильною підпослідовністю. Активація лише частини нейронів з повного представлення недостатньо для доповнення послідовності. Ця проблема та її можливий шлях вирішення обговорюється у висновках до статті.

Г. Часова схожість

Останній експеримент показує ефект часової схожості. Для послідовності довжиною $T = 15$ порівнювалося представлення окремих елементів z_i послідовності. На рис. 3.9б показана матриця схожості i -го елемента з j -тим. Діагональні

елементи показують схожість самих з собою і тому $s(\mathbf{z}_i, \mathbf{z}_i) = 1$. З рисунка видно, що події, які складають одну послідовність кодуються схожим чином, і чим ближче елементи в часі, тим більш вони схожі. Також, елементи далі в послідовності кодуються більш схоже, ніж на початку. Це пов'язано з тим, що з довжиною послідовності росте збудливість певної популяції клітин, яка з більшою ймовірністю починає кодувати всі наступні елементи

Отже, проведені експерименти показали, що дана модель зберігає схожість кодувань для схожих послідовностей (експеримент 1), має чутливість до порядку подій (експеримент 2), має високу схожість для підпослідовностей (експеримент 3) та проявляє часову схожість представлення близьких в часі елементів послідовності (експеримент 4).

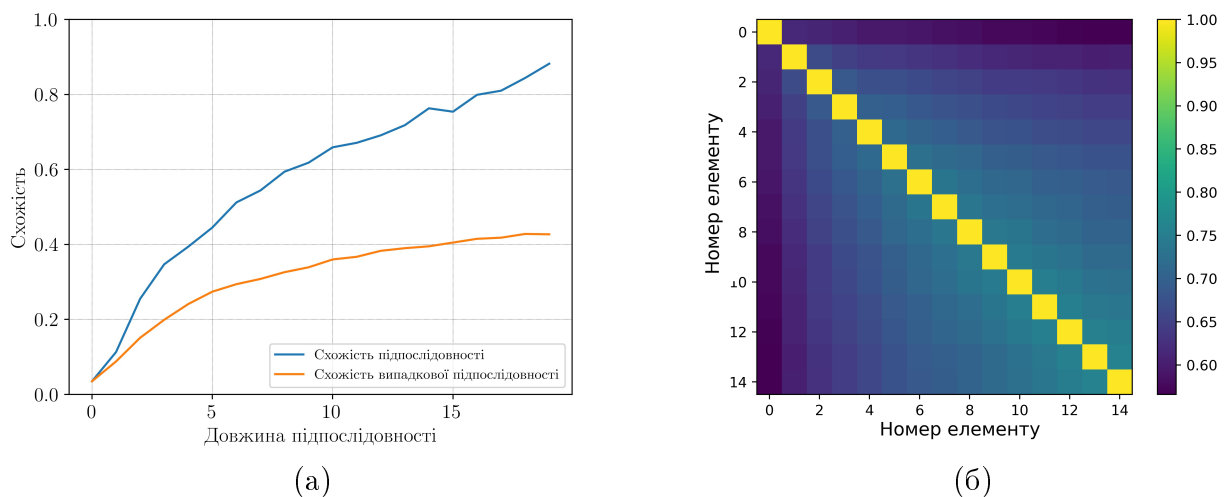


Рис. 3.9: (а) Графік залежності схожості представлень послідовності та її підпослідовності. (б) Перетин популяцій, що кодують індивідуальні елементи для послідовності з 15 елементів

Обговорення до підрозділу

Однією із задач даної роботи було перевірити, чи алокація пам'яті в нейрони з збільшеною здатністю до активації призводить до часової інтеграції послідовності на секундному масштабі часу. Представлені результати якісно повторюють деякі біологічні ефекти, описані в першому розділі, такі як чутливість

до порядку, доповнення послідовності та часова схожість. Ці ефекти спостерігалися для послідовностей на масштабі часу порядку секунд, тому представлені результати вказують на важливість механізму алокації пам'яті навіть на меншому масштабі часу. Підсумовуючи, біологічно подібна модель перетворює вхідні елементи послідовності в розріджений бінарний вектор через модифікований алгоритм випадкової проекції та операції $kWTA()$. Модифікація включає додавання здатності до активації нейронів, що визначається як рухоме середнє активації нейрона. Це призводить до того, що нейрони, що представляли минулі елементи послідовності з більшою ймовірністю будуть представляти і майбутні. Тобто збільшується перетин представлень індивідуальних елементів, як було зображено на рис. 3.7. Друга операція $kWTA()$ вибирає вектор, що представляє всю послідовність. Варто зазначити, що до повного розуміння всіх трьох рівнів за Марром для задачі часової інтеграції ще далеко. В підрозділі представлена постановка задачі та простий алгоритм який відтворює деякі біологічні ефекти. Вважається, що механізм $kWTA$ можуть реалізовувати нейрони інгібітори [62]. Фактор транскрипції CREB регулює збудливість нейрона [81]. Другий $kWTA$ (рівняння 3.5), можливо, реалізують нейрони типу SOM (Somatostatin-expressing interneurons), що виконують латеральну інгібіцію [61]. Для повного відтворення біологічних ефектів потрібно замість неявного механізму $kWTA$ додати шар нейронів інгібіторів на навчати латеральні (автоасоціативні) зв'язки в межах шару, що можливо, призведе до решти ефектів, таких як, стабілізація послідовності та зменшення активності при повторюваності. Також, автоасоціативні зв'язки потрібні для повної реалізації доповнення послідовності. Дана модель показує, що представлення підпослідовності має високий перетин з представленням всієї послідовності, але щоб її видобути потрібно розробляти додаткові механізми та експерименти.

3.5 Висновки

У даній роботі представлено дві моделі нейронних мереж для прогнозування та інтеграції послідовності, що є формами просторово-часової асоціативної пам'яті.

В першій моделі використовується модель нейрона сігма-пай та розріджена активації нейронної мережі. Показано, що така комбінація призводить до високої ємності пам'яті переходів елементів послідовності, що може бути використано для прогнозування послідовності. Дана модель є біологічно подібною, і показано, що мережа з 1000 нейронів та 4000 зв'язків на нейрон має максимальну ємність в 10^5 переходів, що на порядок вище, ніж в подібній моделі НТМ.

Друга модель модифікує алгоритм випадкової проекції вводячи здатність нейрона до активації. Таке нововведення походить від біологічного факту алокації пам'яті в нейрони зі збільшеною здатністю до активації на часовому масштабі годин. Результати даної роботи вказують на можливу важливість алокації навіть на малих масштабах часу порядку секунд, так як обчислювальна модель відтворює деякі біологічні ефекти. Запропонована модель виконує важливу функцію – представлення послідовності, що може бути використано для формування контексту для моделі прогнозування послідовності, для асоціювання та розпізнавання послідовностей.

Підсумовуючи:

1. Застосування моделі нейрона з дендритами до задачі прогнозування послідовності показало високу ємність пам'яті переходів елементів послідовності відносно малої кількості зв'язків завдяки розподіленому кодуванню.

2. Формалізація ідеї алокації пам'яті в нейрони зі збільшеною здатністю до активації якісно відтворює біологічні ефекти представлення послідовностей такі як: збереження схожості, чутливість до порядку, доповнення послідовності та часову схожість.

3. Всі запропоновані моделі та алгоритми є біологічно подібними, тобто не

суперечать відомим уявленням про роботу реальних нейронних мереж. Це важливо для сумісності з майбутніми алгоритмами, розроблених для інших задач, і тому результати даної роботи можуть бути інтегровані в цілісну модель інтелектуального агента.

Результати експериментальних досліджень даного розділу наведено в публікаціях [129, 130]:

1. Osaulenko, V.M Simple model for sequence prediction based On dendritic spatiotemporal integration System Research Information Technologies, 2018, № 4,133-141, DOI:10.20535/SRIT.2308-8893.2018.4.11 (Фахове видання)

2. В. М. Осауленко Модель інтелектуального агента в одновимірному дискретному світі ISSN 1999-9941, “Інформаційні технології та комп’ютерна інженерія”, 2018, № 3,30-36, DOI: <https://doi.org/10.31649/1999-9941-2018-43-3-30-36>. (Фахове видання)

Розділ 4

Застосування моделей на прикладі навчання робота

В останньому розділі дисертації показано як отримані результати з просторово-часової асоціативної пам'яті можуть бути застосовані на прикладі навчання робота. Робототехніка ще знаходиться на ранньому етапі і стикається з трьома основними задачами:

- пошук компактних та ємних джерел енергії,
- розробка нових актуаторів (що рухають частини робота),
- розробка алгоритмів розпізнавання та контролю для роботів.

Перші дві вимагають нових матеріалів та технічних рішень. Третя – пов'язана з необхідністю нової математичної теорії обробки інформації в розподілених системах типу нейронних мереж. З поміж цих трьох, третя є найважчою, так як її вирішення означатиме створення штучного інтелекту, коли робот зможе пізнавати та розуміти взаємозв'язки у світі, приймати рішення і взаємодіяти з середовищем. Ще з 40-х років почався рух зі створення інтелектуальних машин, які б змогли навчатися та виконувати складні задачі як люди. Але ця задача виявилася надзвичайно складною і досі невирішена. Не дивлячись на нещодавній прогрес у сфері глибоких нейронних мереж для розпізнавання образів та навчання з підкріпленням, досі немає чіткого визначення інтелекту та розуміння шляху його відтворення в машині.

Дана робота притримується та розвиває підхід, згідно якого, для створення розумних машин, потрібно спершу зрозуміти принципи роботи біологічних машин. І тоді, ці принципи, формалізовані і перетворенні в математичну абстрактну теорію, можна буде покласти в основу для створення штучних машин. Згідно цього підходу вже довгий час розвиваються brain based devices (BBD, ма-

шини на основі мозку) [93, 37, 44, 113], або ще їх називають NOMAD (neurally organized mobile adaptive devices). На роботах з серії Darwin [152] проводилися чисельні експерименти біологічно правдоподібного навчання. Роботів навчали орієнтуватися в просторі, розпізнавати перешкоди та оминати їх, під'їжджати до "хороших" речей, брати їх та оминати "погані". І все це з використанням біологічно подібних алгоритмів з реалістичною динамікою нейронних мереж.

Існують і інші спроби побудови роботів на основі методів DeepLearning. Проте є кілька причин чому ці методи, поки що, не підходять для задач робототехніки. По перше, вони занадто обчислювально затратні, а мобільні роботи мають обмеження по розмірах, вазі та енергії. По друге, робот має навчатися online, а в deeplearning при появі нових даних чи класів потрібно суттєво перенавчати мережу. По третє, робот має навчатися one shot, тобто вміти розпізнавати нові об'єкти за малої кількості прикладів його зображення. В той же час, біологічні організми здатні до адаптивного та швидкого навчання витрачаючи всього лиш кілька десятків ват. Тому й виникає мотивація вивчати біологічні організми та розвивати напрямок BBD роботів.

4.1 Схеми навчання робота Nao

Розглянемо популярну платформу для дослідження алгоритмів в сфері робототехніки – робота Nao (рис. 4.1). Гуманоїдний робот має висоту 58см та 25 ступенів волі, що дозволяє йому ходити, обстежувати сцену рухами голови, виконувати обмежені дії руками. Стандартна операційна система Naoqi має добре розроблений та документований арі, а також закладено різні готові алгоритми, від ходьби до розпізнавання образів через дві камери на голові. Образи до розпізнавання були попередньо навчені і їх кількість обмежена. Передбачена можливість додавання нового образу до розпізнавання. Спершу, користувач повинен підключити робота до комп'ютера через wi-fi або Ethernet кабель. Далі, новий об'єкт поміщається в поле зору робота і користувач через програмний ін-

терфейс фотографує об'єкт. Тоді користувач повинен вручну виділити контури об'єкта і дати йому назву. Для кращого розпізнавання бажано повторити процедуру для різних сторін об'єкта. Очевидно, що такий спосіб навчання занадто повільний та трудомісткий і потребує вдосконалення.



Рис. 4.1: Робот NAO

Розроблені біологічно подібні моделі просторово-часової асоціативної пам'яті реалізують навчання наближене до біологічних організмів. Тому дані моделі дозволяють навчатися онлайн з малої кількості прикладів, що правда містять недоліки про які буде йтися далі. Просторова асоціативна пам'ять належить до задачі розпізнавання статичних образів, наприклад, об'єктів на зображенні. Часова асоціативна пам'ять належить до задачі розпізнавання рухомих образів, тобто часових послідовностей, наприклад, помах руки вліво.

В роботі запропонована нова схема(рис. 4.2) навчання робота онлайн з використанням отриманих результатів. Загалом, вона виглядає наступним чином. Два сенсорних канали слугують входом, візуальний через камеру та звуковий через мікрофон. З візуального каналу будується карта виразності для трьох каналів(колір, контури, рух), яка комбінується в одну. Виділяється найбільш виразне місце на карті і туди робот спрямовує погляд. Інформація(вектор значень пікселів) з невеликого вікна навколо точки інтересу для кожного каналу через алгоритм випадкової проекції перетворюється у бінарний розріджений вектор, який включає в себе локальні риси об'єкту інтересу. Значення вектору оновлюються з часом для кожного нового погляду робота (тобто для нових рис цього об'єкту). Шар часової інтеграції перетворює послідовність бінарних векторів, що відповідають послідовності виразних рис об'єкта, в один бінарний



Рис. 4.2: Запропонована схема навчання онлайн робота NAO

вектор через алгоритм часової інтеграції, запропонований в третьому розділі. Таким чином в цьому шарі міститься інформація про весь об'єкт або про послідовність об'єктів, якщо робот переводив погляд по різним об'єктам. Звуковий канал також проходить обробку. Може застосовуватися один з існуючих алгоритмів розпізнавання голосу (наприклад, на основі нейронних мереж з модулями LSTM), який переводить звуковий сигналі в слово. Далі, також за алгоритмом випадкової проекції слово(яке є щільним бінарним вектором) перетворюється в розріджений бінарний вектор. Альтернативно, можна вводити текстову назву об'єкта через клавіатуру. Отримані бінарні вектори, два візуальних та один звуковий, асоціюються через алгоритм просторової асоціативної пам'яті, запропонований в розділі два. Виходом схеми є бінарний вектор зі звукового каналу, який може активуватися за відсутності вхідного звуку через асоціативні зв'язки при чому і відбувається розпізнавання.

Загалом, навчання виглядає наступним чином: користувач показує роботу новий об'єкт, рухає ним, привертаючи увагу. Робот центрує погляд на об'єкт та візуальна інформація перетворюється в бінарний вектор. Одночасно користувач

говорить назву об'єкта, яка розпізнається і також перетворюється в бінарний вектор. Між двома векторами створюються асоціативні зв'язки. Якщо потрібно розпізнати не тільки об'єкт, але й рух, наприклад, помах руки вліво, то використовується алгоритм часової інтеграції, запропонований в третьому розділі. Послідовність рис, представлених у вигляді бінарних векторів інтегрується в новий бінарний вектор, який асоціюється із звуковою міткою. В майбутньому, коли робот побачить знову цей об'єкт або рух, через асоціативні зв'язки видобувається відповідна мітка, що і буде виходом алгоритму. В загальному випадку вихід використовується для ініціювання нових рухів роботом, наприклад, синтез тексту назви об'єкту. Для цього використовуються алгоритми генерації послідовності, які згадуються в даній роботі.

Карта виразності (saliency map) вибирає найбільш інформативні риси з візуального потоку. Подібно до того, як це відбувається в живих організмах, візуальна інформація розбивається на модальності: колір, контури, рухи. Далі застосовуються відомі підходи побудови карт виразності [76, 18]. Зокрема, рухи визначаються як різниця між миттєвим та усередненим зображенням через експоненційне згладжування. Вибираючи максимальне значення з різниці отримуємо місце з найбільш виразним рухом. Також отримуються найбільш виразні місця в карті контурів та карті кольорів. Три позиції усереднюються та отримується місце куди робот повертає голову. Вибір малого вікна навколо цього місця подібно до наявності центральної ямки на сітківці ока, яка дає найбільшу роздільну здатність та найкраще виділяє локальні риси об'єкта. Після зміни погляду робота, минуле місце на карті виразності пригнічується, щоб робот знову туди не повертався, що також має відповідний аналог у візуальній системі тварин.

Варто зазначити, що така карта виразності далека від досконалості, і побудова хороших карт є активним науковим напрямком. Проте, навіть така примітивна карта дозволяє послідовно сканувати сцену, подібно до саккад у тварин. Є кілька теорій, які стверджують, що таке сканування необхідне не тільки че-

рез те, що лише мала частина візуального поля має високу роздільну здатність (через ту ж центральну ямку), але і для визначення просторових співвідношень локальних рис об'єкта [63, 118]. Переводячи погляд в інше місце, візуальна кора отримує не тільки нову інформацію з сітківки, але й інформацію з моторних ділянок про напрямок та величину руху. Таким чином, візуальна інформація про локальні риси може інтегруватися з інформацією про її положення відносно попередньої риси або об'єкта в цілому.

В даній схемі застосовується алгоритм випадкової проекції. Раніше вже писалося, що він дозволяє зберігати схожість між схожими вхідними даними, а також є простим способом кодування інформації в бінарні розріджені вектори. Більше того, він використовується в сенсорній систем деяких простих тварин. Проте, в складнішій візуальній системі ваги не є випадковими, а навчаються з досвіду, і як припускається, відтворюють розподіл ймовірностей вхідних стимулів [139]. Тому, в майбутньому випадкову проекцію слід замінити на більш досконалий алгоритм.

Для створення просторових асоціативних зв'язків можна було б використати і інший алгоритм, наприклад, класичну модель Вілшоу. Проте, запропонована модель на основі сігма-пай нейрона дозволяє збільшити ємність пам'яті, а отже кількість можливих образів для розпізнавання. Також, реальні дані такі як візуальні риси сильно корельовані, тому класичні моделі мають вищу похибку. В запропонованій моделі, вибираючи більший порядок збігів (параметр k), можна розрізняти високо корельовані вектори.

При описі задачі розпізнавання для роботи, стає більш зрозумілою необхідність часової інтеграції. Якщо потрібно розпізнати не об'єкт, а рух, наприклад, падіння м'ячика, то потрібно перетворити послідовність візуальних кадрів в одне представлення, яке можна асоціювати з міткою цього руху. При цьому важливо, щоб це представлення мало властивість збереження схожості, подібно як алгоритм випадкової проекції, для забезпечення генералізації. Таким чином, подібні рухи будуть мати подібні представлення, а отже з більшою ймовірністю

будуть правильно розпізнаватися.

Варто зазначити, що до схеми можна додати можливість прогнозування через латеральні зв'язки в шарі одразу після випадкової проекції. Таким чином, після навчання, робот зможе передбачити майбутні рухи, активуючи відповідне представлення в шарі часової інтеграції, який в свою чергу видобуває мітку назви цього руху. Ця інформація може бути використана для початку відповідних моторних команд, наприклад, підставити руку під м'яч, що падає. І тут проявляється вся важливість робити правильні прогнози, так як для успішно виконаної задачі, моторні команди треба починати за багато часу до кінцевої мети.

Дана схема в дещо спрощеному вигляді була реалізована на реальному роботі NAO при стажуванні в лабораторії LORIA у Франції. Було показано її життєздатність (відео демонстрація за посиланням http://bit.do/robot_nao). Скріншот показано на рис.4.3, на якому робот слідує за рухами пальців. На екрані монітора показано зліва направо: візуалізацію бінарних розріджених векторів для асоціації (видно погано), зображення з камери, карта виразності, область інтересу. Очевидно, що така схема набагато краща, ніж підключати робота до комп'ютера, фотографувати, вирізати, називати для кожного об'єкта. Спрощення схеми було в карті виразності, яка не включала модальність контурів, та замість звукового каналу був текстовий. Тобто для перевірки схеми мітка все-таки вводилася через клавіатуру, хоча нічого не забороняє підключити звичайний алгоритм розпізнавання голосу.

На жаль, хоч і запропонована схема краща стандартної схеми навчання, вона містить багато недоліків. По-перше, карта виразності нестабільна, не завжди робот дивиться на найбільш "цікаві" місця. По-друге, точність розпізнавання набагато гірша ніж в класичних згорткових мережах. Дуже слабка генералізація та інваріантність до масштабу та поворотів. Цей недолік пов'язано зі способом кодування бінарних векторів через випадкову проекцію. Потрібно додати навчання ваг та зробити ієрархічну архітектуру подібно до багат шарових ней-

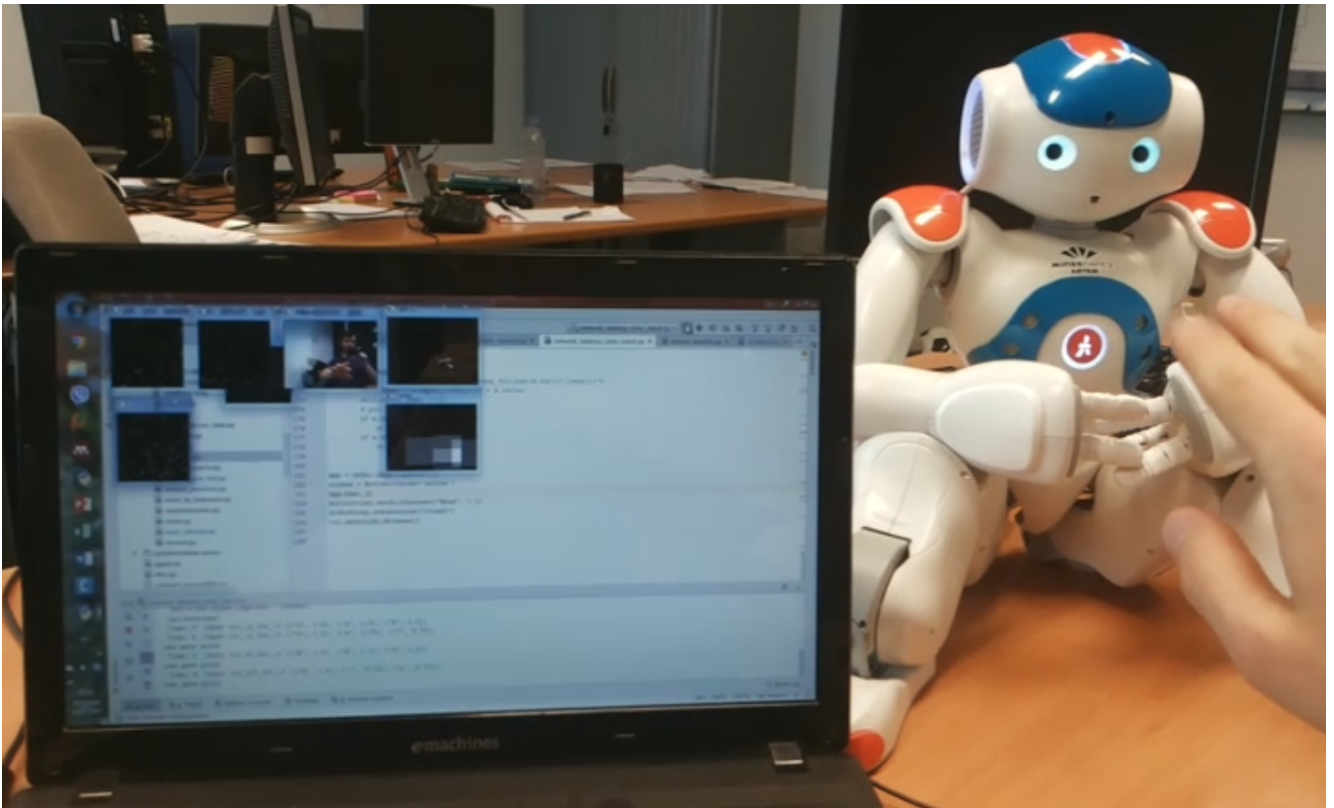


Рис. 4.3: Скріншот з відео демонстрації навчання робота Nao

ронних мереж. Тобто, поки що для розпізнавання образів в робототехніці потрібно поступитися точністю розпізнавання задля навчання online та one shot. По-третє, бракує математичної теорії обробки інформації, яка зможе зв'язати всі складові схеми до купи. Тому що, і в цій роботі, і у подібних, відбувається лише наслідування біологічних механізмів. Є розуміння необхідності тієї чи іншої складової, наприклад, карти виразності, чи саккад, але всі ці складові роз'єднані. І важкість кожної складової підтверджує той факт, що існують цілі наукові напрямки з щорічними конференціями, що досліджують лише певні складові запропонованої схеми (ті ж самі карти виразності, чи асоціативна пам'ять, чи розпізнавання послідовностей).

Варто відзначити, що й методи deeplearning також намагаються вирішити задачу розпізнавання образів в робототехніці. Спеціально створений датасет iCubWorld Transformations Dataset містить фотографії 200 об'єктів з 20 категорії (станом на 2019 рік). Кожний об'єкт фотографується під різними кутами, на різних відстанях, з різним фоном і всього 3600 фото на об'єкт. Така кількість

необхідна для генералізації при навчання в глибоких нейронних мережах, які показують 70-90% точності розпізнавання в залежності від архітектури. Таке навчання важко назвати біологічно подібним, через потребу великої кількості даних та важкості перенавчання при появі нових, проте такі результати значно кращі за біологічно подібні алгоритми, які показують 10-50% точності. Як зазначалося, такі системи розпізнавання мають багато модулів, і якість роботи кожного модуля, яка визначається інженерним рішенням, визначає загальну здатність до розпізнавання. Тому важливо поєднати end-to-end навчання в deep learning з біологічно подібними алгоритмами в загальну теорію обробки інформації в нейронних мережах, чого досі ще нікому не вдалося.

Загалом, отримані результати дали свій маленький внесок в загальну задачу, на вирішення якої підуть далеко ще не одна кандидатська чи докторська.

Результати досліджень даного розділу наведено в публікації [126]:

Osaulenko V.M. On the quest to recreate intelligence . Review of Exploring robotic minds : Actions , symbols , and consciousness as self-organizing dynamic. Cognitive Systems Research, 2018, 51, 56–60. <https://doi.org/10.1016/j.cogsys.2018-04.008> (Іноземне видання, реферується в WebOfScience)

Висновки

У дисертації наведене теоретичне узагальнення і нове вирішення наукової задачі, що виявляється у формалізації біологічних відомостей дендритних обчислень та алокації пам'яті в нейрони зі збільшеною здатністю до активації у моделі просторово-часової асоціативної пам'яті. Також, задача включає проведення аналізу обчислювальних властивостей, можливості практичного застосування, та визначення переваг і недоліків в порівнянні з існуючими підходами.

За підсумками виконаної роботи основні наукові та практичні результати полягають в наступному:

1. Побудована двовірна модель пластичності з різними часовими масштабами сили зв'язку, дозволяє навчати стабільні просторові асоціації (зв'язки), виділяти статистично значимі патерни, та включає сигнал підкріплення. Модель подібна до існуючих підходів, проте включає актуальні біологічні дані взаємодії молекулярних механізмів в синапсах. Робиться припущення про важливість кластеризації АМПА рецепторів для їх стабілізації та підтримки сили зв'язку.
2. Запропонована нова модель асоціативної пам'яті, що використовує нелінійні властивості дендритів та розріджену активацію нейронної мережі призводить до високої ємності асоціативної пам'яті. Ємність пам'яті на основі сігма-пай нейрона обернено пропорційна розрідженості вихідного вектору та комбінаторно розширеного вхідного вектору. Мультиплікативна активація має кращу синаптичну ємність на 40%, але гіршу (на 400%) інформаційну, згідно порівняння з моделлю Вілшоу.
3. Застосування моделі нейрона з дендритами до задачі прогнозування послідовності показало високу ємність пам'яті: наприклад, для розміру мережі $N = 1000$ ємність складає 10^5 порівняно з $\sim 10^4$ для відомої моделі НТМ. Покращення за рахунок мультиплікативної активації, що дозволяє

зменшити кількість зв'язків. Використання часової затримки для контексту спрощує архітектуру мережі, проте накладає обмеження на дальність прогнозування.

4. Побудована модель представлення послідовності в бінарне розріджено-розподілене представлення, якісно відтворює такі біологічні ефекти як, збереження схожості, чутливість до порядку, часову схожість. Теоретичні результати підтверджують гіпотезу про важливість алокації пам'яті в збудливі нейрони навіть на коротких масштабах часу. Робота дає стимул для подальшої експериментальної перевірки гіпотези біологами.

Дані моделі застосовувалися в задачі розпізнавання візуальних образів в робототехніці, в умовах, коли навчальна вибірка мала, навчання онлайн з динамічною кількістю класів. Проте, якість розпізнавання з використанням біологічно подібної асоціативної пам'яті значно гірша за сучасні методи на основі глибокого навчання. Біологічно подібні алгоритми, що реалізують навчання без учителя ще на зародковому етапі і потребують подальших досліджень.

Додатки

1.Список опублікованих праць за темою дисертації

1. Osaulenko, V., Girau, B., Makarenko, O. Henaff, P. Increasing Capacity of Association Memory by Means of Synaptic Clustering. Neural Process. Lett. (2019), pp 1–14 doi:10.1007/s11063-019-10051-7 Особистий внесок: розробка моделі, отримання теоретичних та обчислювальних результатів. (Іноземне видання, реферується в SCOPUS)
2. Osaulenko V.M. On the quest to recreate intelligence . Review of Exploring robotic minds : Actions , symbols , and consciousness as self-organizing dynamic. Cognitive Systems Research, 2018, 51, 56–60. <https://doi.org/10.1016/j.cogsys.2018.04.008> (Іноземне видання, реферується в WebOfScience)
3. Osaulenko V.M Simple model for sequence prediction based on dendritic spatiotemporal integration System Research & Information Technologies, 2018, № 4, 133-141, DOI:10.20535/SRIT.2308-8893.2018.4.11 (Фахове видання, індексується наукометричними базами DOAJ, Index Copernicus, Google Scholar, PBN)
4. В. М. Осауленко Модель інтелектуального агента в одномірному дискретному світі ISSN 1999-9941, “Інформаційні технології та комп’ютерна інженерія”, 2018, № 3, 30-36, DOI: <https://doi.org/10.31649/1999-9941-2018-43-3-30-36>. (Фахове видання, індексується Google Scholar)
5. Осауленко В.М. Двовимірна модель навчання у спайкових нейронних мережах з гомеостазом та навчанням з підкріпленням, System Research & Information Technologies, 2017, № 2, 130-140, DOI: 10.20535/SRIT.2308-8893.

- 2017.1.12 (Фахове видання, індексується наукометричними базами DOAJ, Index Copernicus, Google Scholar, PBN)
6. Osaulenko V.M. Testing simple neuron models with dendrites for sparse binary image representation, Штучний інтелект, 2017, № 2, С. 101-108 (Фахове видання, індексується Google Scholar)
 7. Makarenko O.S., Osaulenko V.M. Application of cellular automates in some models of artificial intelligence. Конференція IEEE SAIC-2018, 2018, 1-4. 10.1109/SAIC.2018.8516837
 8. Osaulenko V.M. Learning of higher-order correlations improves capacity of an association memory. Nonlinear analysis and applications 4rd International Conference of memory of corresponding member of National Academy of Sciences of Ukraine Valeriy Sergeevich MELNIK (1952-2007) NTUU KPI, Kyiv, Ukraine, April 04-06, 2018, p.76.
 9. Osaulenko V.M. Model for sequence prediction based on dendritic spatio-temporal integration. 20-th International conference on System Analysis and Information Technology SAIT 2018, May 21-23, 2018, pp.36-37. Institute for Applied System Analysis of National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute Kyiv, Ukraine
 10. Osaulenko V.M. Dendritic nonlinearity provides high capacity sequence memory FENS FORUM of neuroscience 2018, p.368, Berlin
 11. Osaulenko V.M. New ideas to test for artificial neural networks inspired from neuroscience Міжнародна наукова конференція «Сучасна інформатика: проблеми, досягнення та перспективи розвитку» Інституту кібернетики імені В.М.Глушкова, 13-15 грудня 2017, с.82-83
 12. Осауленко В.М. Тестування простих моделей з дендритами для розрізженого бінарного представлення зображень. XVII Міжнародна науково-

технічна конференція Штучний інтелект та інтелектуальні системи. Artificial intelligence and intellegent systems (aiis'2017), Київ 2017, с. 101-108

13. Осауленко В.М. На шляху до створення інтелекту Philosophed 2015 міжнародна наукова конференція, ст. 47-49, Київ 2015.

2. Довідка про впровадження в навчальний процес КПІ ім. Ігоря Сікорського

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ * МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ



ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ
НАЦІОНАЛЬНОГО ТЕХНІЧНОГО УНІВЕРСИТЕТУ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
ім. Ігоря Сікорського”

№ 62/4040
“14” 05 2019



Затверджую

Заступник Директора ІПСА НТУУ «КПІ»

з навчально-наукової роботи,

д.т.н., професор

В.Д. Романенко

**про впровадження результатів дисертаційної роботи
Осауленка В'ячеслава Миколайовича
у навчальний процес ІПСА НТУУ «КПІ ім. І. Сікорського»**

У процесі виконання дисертаційної роботи Осауленко В'ячеслав Миколайович отримав ряд нових наукових результатів стосовно асоціативної пам'яті для статичних та динамічних даних. Отримані результати впроваджено у навчальний процес кафедри Математичних методів системного аналізу Інституту прикладного системного аналізу НТУУ «КПІ ім. І. Сікорського».


Зокрема, побудовано модель пам'яті на основі сігма-пай нейрона, що використовує мультиплікативну активаційну функцію, отримано аналітичні вирази ємності пам'яті та порівняно з модифікованою моделлю Вілшоу. Навчання асоціації було розширено для динамічних даних, що дозволяє робити передбачення елементів послідовності. Досліджено модель нейронної мережі з алокацією пам'яті в збудливі нерони, і показано, що обчислювальні ефекти якісно відтворюють біологічні. Це свідчить про важливість збудливості нейрону для представлення послідовностей в нейронній мережі.

Результати досліджень увійшли в практичні роботи для дисципліни «Нейронні мережі» з навчального плану Інституту прикладного системного аналізу. Використовується детальний опис моделей асоціативної пам'яті з програмною реалізацією у вигляді jupyter notebook.

Професор ІПСА НТУУ «КПІ»,
д.т.н.

 В.Я. Данилов

Професор ІПСА НТУУ «КПІ»,
д.ф.м.н.

 О.С. Макаренко

03056, Київ-56, проспект Перемоги 37. Тел. (044) 204-96-47, 204-8706. Тел./факс: (044) 236-3987
e-mail: mmsa@mmsa.ntu-kpi.kiev.ua

3. Акт впровадження в роботу лабораторії Loria м.Нансі, Франція



Campus ARTEM
92, rue du sergent Blandan -
CS 14234
54042 Nancy cedex
FRANCE

Nancy, 25 September 2017

Act of implementation

To whom it may concern,

I hereby, Patrick Hénaff, Professor at University of Lorraine (France), certify that Osaulenko Viacheslav, PhD student from Kiev Polytechnic Institute, worked in our LORIA laboratory under our supervision (me and Prof. B. Girau) from January 2017 to July 2017.

This document verifies that results of research provided by Osaulenko Viacheslav were useful for our work and found the practical implementation. Particularly, algorithms developed for storing association memories introduces new ideas for neural implementation of Hebbian learning. Also, these results are considered to be implemented in humanoid robotics, specifically, in real time processing of visual input and creating an association between objects.

Results of this research were included into the final thesis of the dissertation.

Professeur Patrick Hénaff,
Ecole des Mines de Nancy
Université de Lorraine
France

Бібліографія

- [1] L. F. Abbott. Lapticque's introduction of the integrate-and-fire model neuron (1907), 1999.
- [2] L. F. Abbott and Y. Arian. Storage capacity of generalized networks, 1987.
- [3] L. F. Abbott and S. B. Nelson. Synaptic plasticity: taming the beast. *Nature neuroscience*, 3(november):1178–1183, 2000.
- [4] E. J. Agnes, R. E. Jr., and L. G. Brunnet. *Associative Memory in Neuronal Networks of Spiking Neurons: Architecture and Storage Analysis*. 2012.
- [5] S. Ahmad and J. Hawkins. Properties of Sparse Distributed Representations and their Application to Hierarchical Temporal Memory. *CoRR*, abs/1503.0:1503.07469, 2015.
- [6] S. Ahmad and J. Hawkins. How do neurons operate on sparse distributed representations? A mathematical theory of sparsity, neurons and active dendrites. *arXiv*, pages arXiv:1601.00720 [q-bio.NC], jan 2016.
- [7] Z. Alom, B. V. Essen, A. T. Moody, D. P. Widemann, and T. M. Taha. Convolutional Sparse Coding on Neurosynaptic Cognitive System. *IEEE*, pages 3609–3616, 2017.
- [8] D. J. Amit, H. Gutfreund, and H. Sompolinsky. Information storage in neural networks with low levels of activity. *Physical Review A*, 1987.
- [9] P. Baldi. Number of Stable Points for Spin-Glasses and Neural Networks of Higher Orders. *Physical Review Letters*, 58(9), 1987.
- [10] H. Barlow. Redundancy reduction revisited. *Network: Computation in Neural Systems*, 12(3):241–253, 2001.

- [11] H. Barlow, T. P. Kaushal, and G. J. Mitchison. Finding minimum entropy codes. *Neural Comp.*, 1:412–423, 1989.
- [12] A. J. Bell and T. Sejnowski. The “independent components” of natural scenes are edge filters. *Vision Research*, 37(23):3327–3338, 1997.
- [13] S. Bengio, Y. Bengio, and J. Cloutier. On the search for new learning rules for ANNs. *Neural Processing Letters*, 2(4):26–30, 1995.
- [14] U. S. Bhalla. Synaptic input sequence discrimination on behavioral timescales mediated by reaction-diffusion chemistry in dendrites. *eLife*, 6:1–24, 2017.
- [15] U. U. S. Bhalla. Dendrites, Deep Learning, and Sequences in the Hippocampus. *Hippocampus*, 2014(6), 2017.
- [16] E. L. Bienenstock, L. N. Cooper, and P. W. Munro. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience*, 2(1):32–48, 1982.
- [17] K. C. Bittner, A. D. Milstein, C. Grienberger, S. Romani, and J. C. Magee. Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science*, 357(6355):1033–1036, 2017.
- [18] A. Borji and L. Itti. State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):185–207, 2013.
- [19] J. Bose, S. B. Furber, and J. L. Shapiro. An associative memory for the on-line recognition and prediction of temporal sequences. *Proceedings of the International Joint Conference on Neural Networks*, 2:1223–1228, 2005.
- [20] T. Branco. Computing Temporal Sequence with Dendrites. *chapter*, 11:245–257, 2014.

- [21] T. Branco, B. A. Clark, and M. Häusser. Dendritic discrimination of temporal input sequences in cortical neurons. *Science*, 329(5999):1671–1675, 2010.
- [22] T. Branco and M. Häusser. The single dendritic branch as a fundamental functional unit in the nervous system. *Current Opinion in Neurobiology*, 20(4):494–502, 2010.
- [23] J. Brea and W. Gerstner. Does computational neuroscience need new synaptic learning paradigms? *Current Opinion in Behavioral Sciences*, pages 1–6, 2016.
- [24] R. Brette and W. Gerstner. Neuronal Activity. *Journal of Neurophysiology*, 94:3637–3642, 2005.
- [25] D. J. Cai, D. Aharoni, T. Shuman, J. Shobe, J. Biane, W. Song, B. Wei, M. Veshkini, M. La-Vu, J. Lou, S. E. Flores, I. Kim, Y. Sano, M. Zhou, K. Baumgaertel, A. Lavi, M. Kamata, M. Tuszynski, M. Mayford, P. Golshani, and A. J. Silva. A shared neural ensemble links distinct contextual memories encoded close in time. *Nature*, 534(7605):115–118, 2016.
- [26] H. H. Chen, G. Z. Sun, T. Maxwell, H. Y. Lee, and C. L. Giles. MACIUNE LEARNING USING A HIGHER ORDER CORRELATION NETWORK Y.C. LEE and Gary DOOLEN. *Physica A: Statistical Mechanics and its Applications*, page 1986, 1986.
- [27] M. Chistiakova, N. M. Bannon, M. Bazhenov, and M. Volgushev. Heterosynaptic Plasticity: Multiple Mechanisms and Multiple Roles. *The Neuroscientist*, 20(5):483–498, oct 2014.
- [28] A. Clark. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(May 2017):181–253, 2013.

- [29] Y. Cui, S. Ahmad, and J. Hawkins. The HTM Spatial Pooler: a neocortical algorithm for online sparse distributed coding. *bioRxiv*, 11(November):0–15, 2016.
- [30] Y. Cui, S. Ahmad, J. Hawkins, C. Surpur, S. Ahmad, J. Hawkins, C. Surpur, S. Ahmad, and J. Hawkins. Continuous online sequence learning with an unsupervised neural network model. *arXiv:1512.05463 [cs, q-bio]*, 698:672–698, 2015.
- [31] Y. Cui, C. Surpur, S. Ahmad, and J. Hawkins. A comparative study of HTM and other neural network models for online sequence learning with streaming data. *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 1530–1538, 2016.
- [32] M. D. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman;, 1982.
- [33] S. Dasgupta, C. F. Stevens, and S. Navlakha. A neural algorithm for a fundamental computing problem. *Science*, 358(6364):793–796, 2017.
- [34] L. Davachi and S. DuBrow. How the hippocampus preserves order: The role of prediction and context. *Trends in Cognitive Sciences*, 19(2):92–99, 2015.
- [35] R. J. Douglas, K. A. C. Martin, and D. Whitteridge. Microcircuit for Neocortex. *Neural Computation*, 1989.
- [36] K. Du, Y.-W. Wu, R. Lindroos, Y. Liu, B. Rózsa, G. Katona, J. B. Ding, and J. H. Kotaleski. Cell-type-specific inhibition of the dendritic plateau potential in striatal spiny projection neurons. *Proceedings of the National Academy of Sciences*, 114(36):E7612–E7621, 2017.
- [37] G. M. Edelman. Learning in and from Brain-Based Devices. *Science*, 318(5853):1103–1105, 2007.

- [38] A. K. Engel, P. König, A. K. Kreiter, T. B. Schillen, and W. Singer. Temporal coding in the visual cortex: new vistas on integration in the nervous system. *Trends in Neurosciences*, 15(6):218–226, 1992.
- [39] J. a. Esteban, S.-H. Shi, C. Wilson, M. Nuriya, R. L. Huganir, and R. Malinow. PKA phosphorylation of AMPA receptor subunits controls synaptic trafficking underlying plasticity. *Nature neuroscience*, 6(2):136–143, 2003.
- [40] C. W. C. Eurich, H. Schwegler, and R. Woesler. Coarse coding: applications to the visual system of salamanders. *Biological Cybernetics*, 47:41–47, 1997.
- [41] M. Fauth and C. Tetzlaff. Opposing Effects of Neuronal Activity on Structural Plasticity. *Frontiers in Neuroanatomy*, 10(June):1–18, 2016.
- [42] D. J. Felleman and D. C. Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1(1):1–47, 1991.
- [43] D. J. Field and B. Olshausen. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 1996.
- [44] J. Fleischer and G. M. Edelman. Brain-based devices. *IEEE Robotics & Automation Magazine*, 16(3):33–41, 2009.
- [45] P. Földiák. Forming sparse representations by local anti-Hebbian learning. *Biological Cybernetics*, 64(2):165–170, 1990.
- [46] A. C. Frank, S. Huang, M. Zhou, A. Gdalyahu, G. Kastellakis, T. K. Silva, E. Lu, X. Wen, P. Poirazi, J. T. Trachtenberg, and A. J. Silva. Hotspots of dendritic spine turnover facilitate clustered spine addition and learning and memory. *Nature Communications*, 9(1):422, 2018.
- [47] N. Frémaux and W. Gerstner. Neuromodulated Spike-Timing-Dependent Plasticity and Theory of Three-Factor Learning Rules. *Frontiers in Neural Circuits*, 9(85):85, 2016.

- [48] K. J. Friston and C. D. Frith. Active inference, Communication and hermeneutics. *Cortex*, 68(Kelso 2012):129–143, 2015.
- [49] K. Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4):193–202, 1980.
- [50] a. R. Gardner-Medwin. The recall of events through the learning of associations between their parts., 1976.
- [51] J. P. Gavornik and M. F. Bear. Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nature Neuroscience*, 17(5):732–737, 2014.
- [52] A. P. Georgopoulos, A. B. Schwartz, and R. E. Kettner. Neuronal population coding of movement direction. *Science (New York, N.Y.)*, 233(4771):1416–1419, 1986.
- [53] C. L. Giles and T. Maxwell. Learning, invariance, and generalization in high-order neural networks. *Applied optics*, 26:4972–4978, 1987.
- [54] J. Gjorgjieva, C. Clopath, J. Audet, and J.-P. Pfister. A triplet spike-timing-dependent plasticity model generalizes the Bienenstock-Cooper-Munro rule to higher-order spatiotemporal correlations. *Proceedings of the National Academy of Sciences of the United States of America*, 108(48):19383–8, 2011.
- [55] A. Govindarajan, R. J. Kelleher, S. Tonegawa, and A. Govindarajan Raymond J. Tonegawa, Susumu. A clustered plasticity model of long-term memory engrams. *Nature reviews. Neuroscience*, 7(July):575–583, 2006.
- [56] B. Graham and D. Willshaw. Capacity and information efficiency of the associative net. *Network: Computation in Neural Systems*, 8(1):35–54, 1997.

- [57] V. I. Gritsenko, D. A. Rachkovskij, A. A. Frolov, R. Gayler, D. Kleyko, and E. Osipov. Neural distributed autoassociative memories: A survey. *Cybernetics and Computer Engineering*. *Cybernetics and Computer Engineering*, 2(188):5–35, 2017.
- [58] G. G.Turrigiano, S. B.Nelson, G. G. Turrigiano, S. B. Nelson, G. G.Turrigiano, and S. B.Nelson. Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience*, 5(2):97–107, 2004.
- [59] S. J. Guzman, A. Schlögl, M. Frotscher, and P. Jonas. Synaptic mechanisms of pattern completion in the hippocampal CA3 network. *Science*, 353(6304):109–120, 2016.
- [60] K. D. Harris and T. D. Mrsic-Flogel. Cortical connectivity and sensory coding. *Nature*, 503(7474):51–8, 2013.
- [61] K. D. Harris and G. M. G. Shepherd. The neocortical circuit: themes and variations. *Nature Neuroscience*, 18(2):170–181, 2015.
- [62] J. Hawkins and S. Ahmad. Why Neurons Have Thousands of Synapses, a Theory of Sequence Memory in Neocortex. *Frontiers in Neural Circuits*, 10(March):1–20, 2016.
- [63] J. Hawkins, S. Ahmad, and Y. Cui. Why Does the Neocortex Have Layers and Columns, A Theory of Learning the 3D Structure of the World. *bioRxiv*, July 12(October):0–15, 2017.
- [64] HBP. HBP Neurorobotics Platform, 2018.
- [65] K. He, M. Huertas, S. Z. Hong, X. X. Tie, J. W. Hell, H. Shouval, and A. Kirkwood. Distinct Eligibility Traces for LTP and LTD in Cortical Synapses. *Neuron*, 88(3):528–538, 2015.

- [66] K. He, X. Zhang, S. Ren, and J. Sun. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9):1904–1916, 2015.
- [67] D. O. Hebb. *The Organization of Behavior*. 1949.
- [68] G. E. Hinton, J. L. McClelland, and D. E. Rumelhart. Distributed representations. *Parallel Distributed Processing*, pages 77–109, 1986.
- [69] T. Hofmann and J. M. Buhmann. Competitive Learning algorithms for robust vector quantization. *IEEE Transactions on Signal Processing*, 46(6):1665–1675, 1990.
- [70] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982.
- [71] D. H. Hubel. Evolution of ideas on the primary visual cortex, 1955-1978: A biased historical account - Nobel lecture, 8 December 1981. *Bioscience Reports*, 2(7):435–469, 1982.
- [72] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat’s striate cortex. *The Journal of Physiology*, 1959.
- [73] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 1968.
- [74] R. Hunter, S. Cobb, and B. P. Graham. *Improving Associative Memory in a Network of Spiking Neurons Russell*. 2008.
- [75] K. Ibata, Q. Sun, and G. G. Turrigiano. Rapid Synaptic Scaling Induced by Changes in Postsynaptic Firing. *Neuron*, 57(6):819–826, 2008.
- [76] L. Itti, C. Koch, L. Itti, C. Koch, and C. Koch. Computational modelling of visual attention. *Nature reviews. Neuroscience*, 2(March):194–203, 2001.

- [77] E. M. Izhikevich. Which model to use for cortical spiking neurons? *IEEE Transactions on Neural Networks*, 15(5):1063–1070, 2004.
- [78] E. M. Izhikevich. Polychronization: computation with spikes. *Neural computation*, 18(2):245–82, 2006.
- [79] E. M. Izhikevich. Dynamical Systems in Neuroscience:. *Dynamical Systems*, 25(1):1–383, 2007.
- [80] H. Jia, N. L. Rochefort, X. Chen, and A. Konnerth. Dendritic organization of sensory input to cortical neurons in vivo. *Nature*, 464(7293):1307–1312, 2010.
- [81] S. A. Josselyn and P. W. Frankland. Memory Allocation : Mechanisms and Function. *Annual Review of Neuroscience*, (April):389–413, 2018.
- [82] P. Kanerva. Sparse Distributed Memory. *MIT Press*, 1988.
- [83] P. Kanerva. Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors. *Cognitive Computation*, 1(2):139–159, 2009.
- [84] G. Kastellakis, D. J. Cai, S. C. Mednick, A. J. Silva, and P. Poirazi. Synaptic clustering within dendrites: An emerging theory of memory formation. *Progress in Neurobiology*, 126:19–35, 2015.
- [85] G. Kastellakis, A. J. Silva, and P. Poirazi. Linking Memories across Time via Neuronal and Dendritic Overlaps in Model Neurons with Active Dendrites. *Cell Reports*, 17(6):1491–1504, 2016.
- [86] D. Kersten, P. Mamassian, and A. Yuille. Object Perception as Bayesian Inference. *Annual Review of Psychology*, 55(1):271–304, 2004.
- [87] A. J. King, S. Teki, and B. D. Willmore. Recent advances in understanding the auditory cortex. *F1000Research*, 7(0):1555, 2018.

- [88] D. C. Knill and A. Pouget. The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12):712–719, 2004.
- [89] A. Knoblauch, G. Palm, and F. T. Sommer. Memory Capacities for Synaptic and Structural Plasticity. *Neural Computation*, 22(2):289–341, 2010.
- [90] A. Knoblauch and F. T. Sommer. Structural Plasticity, Effectual Connectivity, and Memory in Cortex. *Frontiers in Neuroanatomy*, 10(June):1–20, 2016.
- [91] C. Koch, T. Poggio, and V. Torre. Nonlinear interactions in a dendritic tree: localization, timing, and role in information processing. *Proceedings Of The National Academy Of Sciences Of The United States Of America-Biological Sciences*, 80(9):2799–2802, may 1983.
- [92] B. Kosko. Bidirectional Associative Memories. *IEEE Transactions on Systems, Man and Cybernetics*, 18(1):49–60, 1988.
- [93] J. L. Krichmar and G. M. Edelman. Machine Psychology: Autonomous Behavior, Perceptual Categorization and Conditioning in a Brain-based Device. *Cerebral Cortex*, 12(8):818–830, 2002.
- [94] J. L. Krichmar, A. K. Seth, D. A. Nitz, J. G. Fleischer, and G. M. Edelman. Spatial Navigation and Causal Analysis in a Brain-Based Device Modeling Cortical–Hippocampal Interactions. *Neuroinformatics*, 5(1):11–34, 2005.
- [95] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. *Advances In Neural Information Processing Systems*, pages 1–9, 2012.
- [96] D. Krotov and J. J. Hopfield. Dense Associative Memory for Pattern Recognition. *Nips*, (Nips), 2016.

- [97] S. H. Lane, D. A. Handelman, and J. J. Gelfand. Theory and Development of Higher-Order CMAC Neural Networks. *IEEE Control Systems*, 12(2):23–30, 1992.
- [98] M. E. Larkum and T. Nevian. Synaptic clustering by dendritic signalling mechanisms. *Current Opinion in Neurobiology*, 18(3):321–331, 2008.
- [99] S.-J. R. Lee, Y. Escobedo-Lozoya, E. M. Szatmari, and R. Yasuda. Activation of CaMKII in single dendritic spines during long-term potentiation. *Nature*, 458(7236):299–304, 2009.
- [100] T. S. Lee and D. Mumford. Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America*, 20(7):1434–48, 2003.
- [101] A. H. Leighton and C. Lohmann. The Wiring of Developing Sensory Circuits—From Patterned Spontaneous Activity to Synaptic Plasticity Mechanisms. *Frontiers in Neural Circuits*, 10(September):1–13, 2016.
- [102] N. Li, T.-W. Chen, Z. V. Guo, C. R. Gerfen, and K. Svoboda. A motor cortex circuit for motor planning and movement. *Nature*, 519(7541):51–56, 2015.
- [103] J. E. Lisman. The Challenge of Understanding the Brain: Where We Stand in 2015. *Neuron*, 86(4):864–882, 2015.
- [104] M. London and M. Häusser. Dendritic Computation. *Annual Review of Neuroscience*, 28(1):503–532, 2005.
- [105] A. Luczak and J. N. Maclean. Default activity patterns at the neocortical microcircuit level. *Frontiers in integrative neuroscience*, 6(June):30, 2012.
- [106] A. Luczak, B. L. McNaughton, and K. D. Harris. Packet-based communication in the cortex. *Nature reviews. Neuroscience*, 16(12):745–755, 2015.
- [107] W. Maass. Networks of spiking neurons: The third generation of neural network models. *Neural Networks*, 10(9):1659–1671, dec 1997.

- [108] M. Mahoney. *Data Compression Explained*. 2013.
- [109] O. S. Makarenko and V. M. Osaulenko. Application of Cellular Automates in Some Models of Artificial Intelligence. *IEEE First International Conference on System Analysis & Intelligent Computing (SAIC)*, pages (pp. 1–4). IEEE., 2018.
- [110] H. Markram, J. Lubke, M. Frotscher, and B. Sakmann. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, 275(5297):213–215, 1997.
- [111] H. Markram, E. Muller, Et.al., S. Ramaswamy, M. W. W. Reimann, M. Abdellah, C. A. A. Sanchez, A. Ailamaki, L. Alonso-Nanclares, N. Antille, S. Arsever, G. A. A. A. Kahou, T. K. K. Berger, A. Bilgili, N. Buncic, A. Chalimourda, G. Chindemi, J.-D. D. Courcol, F. Delalondre, V. Delattre, S. Druckmann, R. Dumusc, J. Dynes, S. Eilemann, E. Gal, M. E. E. Gevaert, J.-P. P. Ghobril, A. Gidon, J. W. W. Graham, A. Gupta, V. Haenel, E. Hay, T. Heinis, J. B. B. Hernando, M. Hines, L. Kanari, D. Keller, J. Kenyon, G. Khazen, Y. Kim, J. G. G. King, Z. Kisvarday, P. Kumbhar, S. Lasserre, J. V. Le Bé, B. R. R. C. Magalhães, A. Merchán-Pérez, J. Meystre, B. R. R. Morrice, J. Muller, A. Muñoz-Céspedes, S. Muralidhar, K. Muthurasa, D. Nachbaur, T. H. H. Newton, M. Nolte, A. Ovcharenko, J. Palacios, L. Pastor, R. Perin, R. Ranjan, I. Riachi, J.-R. R. Rodríguez, J. L. L. Riquelme, C. Rössert, K. Sfyarakis, Y. Shi, J. C. C. Shillcock, G. Silberberg, R. Silva, F. Tauheed, M. Telefont, M. Toledo-Rodriguez, T. Tränkler, W. Van Geit, J. V. V. Díaz, R. Walker, Y. Wang, S. M. M. Zaninetta, J. Defelipe, S. L. L. Hill, I. Segev, F. Schürmann, J.-V. Le Bé, B. R. R. C. Magalhães, A. Merchán-Pérez, J. Meystre, B. R. R. Morrice, J. Muller, A. Muñoz-Céspedes, S. Muralidhar, K. Muthurasa, D. Nachbaur, T. H. H. Newton, M. Nolte, A. Ovcharenko, J. Palacios, L. Pastor, R. Perin, R. Ranjan, I. Riachi, J.-R. R. Rodríguez,

- J. L. L. Riquelme, C. Rössert, K. Sfyrakis, Y. Shi, J. C. C. Shillcock, G. Silberberg, R. Silva, F. Tauheed, M. Telefont, M. Toledo-Rodriguez, T. Tränkler, W. Van Geit, J. V. V. Díaz, R. Walker, Y. Wang, S. M. M. Zaninetta, J. Defelipe, S. L. L. Hill, I. Segev, and F. Schürmann. Reconstruction and Simulation of Neocortical Microcircuitry. *Cell*, 163(2):456–492, 2015.
- [112] W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 1943.
- [113] J. L. McKinstry and G. M. Edelman. Temporal sequence learning in winner-take-all networks of spiking neurons demonstrated in a brain-based device. *Frontiers in Neurorobotics*, 7(JUN):1–10, 2013.
- [114] B. W. Mel. The clusteron: Toward a simple abstraction for a complex neuron. *Nips*, pages 35–42, 1991.
- [115] B. W. Mel. NMDA-Based Pattern Discrimination in a Modeled Cortical Neuron. *Neural Computation*, 4:502–517, 1992.
- [116] B. W. Mel, J. Schiller, and P. Poirazi. Synaptic plasticity in dendrites: complications and coping strategies. *Current Opinion in Neurobiology*, 43:177–186, 2017.
- [117] M. Minsky and S. Papert. Perceptrons. In *Perceptrons: An Introduction to Computational Geometry*. 1969.
- [118] M. B. Mirza, R. A. Adams, C. D. Mathys, and K. J. Friston. Scene Construction, Visual Foraging, and Active Inference. *Frontiers in Computational Neuroscience*, 10(June), 2016.
- [119] V. B. Mountcastle. The columnar organization of the neocortex. *Brain*, 120(4):701–722, 1997.

- [120] H. Mustovic, K. Scheffler, F. Di Salle, F. Esposito, J. G. Neuhoff, J. Hennig, and E. Seifritz. Temporal integration of sequential auditory events: Silent period in sound pattern activates human planum temporale. *NeuroImage*, 20(1):429–434, 2003.
- [121] P. Mutalik. Temporal Sequences in the Brain. *technical report*, (April), 2018.
- [122] L. Nadel, A. Hupbach, R. Gomez, and K. Newman-Smith. Memory formation, consolidation and transformation. *Neuroscience & Biobehavioral Reviews*, 36(7):1640–1645, 2012.
- [123] S.-p. Networks. Chapter 3 Higher-Order Neural Networks. *Neural Networks*, pages 56–79, 1987.
- [124] E. Oja. A simplified neuron model as a principal component analyzer. *Journal of mathematical biology*, 15(3):267–73, 1982.
- [125] B. A. Olshausen and D. J. Field. Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14(4):481–487, 2004.
- [126] V. Osaenko. On the quest to recreate intelligence . Review of Exploring robotic minds : Actions , symbols , and consciousness as self-organizing dynamic. *Cognitive Systems Research*, 51:56–60, oct 2018.
- [127] V. Osaenko, B. Girau, O. Makarenko, and P. Henaff. Increasing Capacity of Association Memory by Means of Synaptic Clustering. *Neural Processing Letters*, 2019.
- [128] V. M. Osaenko. Two dimensional model of learning in spiking neural networks with homeostasis and reward. *System research and information technologies*, 0(2):130–140, 2017.

- [129] V. M. Osaenko. Модель інтелектуального агента в одномірному дискретному світі. *“ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА КОМП’ЮТЕРНА ІНЖЕНЕРІЯ*, 3:30–36, 2018.
- [130] V. M. Osaenko. Simple model for sequence prediction based on dendritic spatiotemporal integration. *System research and information technologies*, 0(4):133–141, 2018.
- [131] G. Palm. Neural associative memories and sparse coding. *Neural Networks*, 37:165–171, 2013.
- [132] P. Peretto and J. J. Niez. Long term memory storage capacity of multi-connected neural networks. *Biological Cybernetics*, 54(1):53–63, 1986.
- [133] J.-P. Pfister and W. Gerstner. Triplets of Spikes in a Model of Spike Timing-Dependent Plasticity. *The Journal of Neuroscience*, 26(38):9673–9682, 2006.
- [134] T. A. Plate. Randomly connected sigma-pi neurons can form associative memories. *Draft*, (1969), 1998.
- [135] T. A. Plate. Distributed representations. *Encyclopedia of Cognitive Science*, (59):77–109, 2012.
- [136] P. Poirazi, T. Brannon, and B. W. Mel. Pyramidal Neuron as Two-Layered Neural Network. *Neuron*, 37:989–999, 2003.
- [137] P. Poirazi and B. Mel. Memory Capacity of Linear vs . Nonlinear Models of Dendritic Integration. *Advances in neural information processing systems*, (November):157–163, 2000.
- [138] P. Poirazi and B. W. Mel. Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron*, 29(3):779–796, 2001.

- [139] D. Purves, B. B. Monson, J. Sundararajan, and W. T. Wojtach. How biological vision succeeds in the physical world. *Proceedings of the National Academy of Sciences of the United States of America*, 111(13):4750–5, 2014.
- [140] R. Q. Quiroga, G. Kreiman, C. Koch, and I. Fried. Sparse but not ‘Grandmother-cell’ coding in the medial temporal lobe. *Trends in Cognitive Sciences*, 12(3):87–91, 2008.
- [141] D. A. Rachkovskij. Estimation of Vectors Similarity by Their Randomized Binary Projections. *Cybernetics and Systems Analysis*, 51(5):808–818, 2015.
- [142] D. A. Rachkovskij and V. I. Gritsenko. *Розподілене подання векторних даних на основі випадкових проєкцій*. Київ: Інтерсервіс, . ISBN 978-617-696-837-5, 2018.
- [143] D. A. Rachkovskij and E. M. Kussul. Binding and Normalization of Binary Sparse Distributed Representations by Context-Dependent Thinning. *Neural Computation*, 13(2):411–452, 2001.
- [144] D. A. Rachkovskij, E. M. Kussul, and T. N. Baidyk. Building a world model with structure-sensitive sparse binary distributed representations. *Biologically Inspired Cognitive Architectures*, 3:64–86, 2013.
- [145] R. P. N. Rao and T. Sejnowski. Predictive Sequence Learning in Recurrent Neocortical Circuits. *Advances in Neural Information Processing Systems 12*, 70:164–170, 2000.
- [146] R. L. Redondo and R. G. M. Morris. Making memories last: the synaptic tagging and capture hypothesis. *Nature reviews. Neuroscience*, 12(1):17–30, 2011.

- [147] T. Rogerson, D. J. Cai, A. Frank, Y. Sano, J. Shobe, M. F. Lopez-Aranda, and A. J. Silva. Synaptic tagging during memory allocation. *Nature reviews. Neuroscience*, 15(3):157–69, 2014.
- [148] E. T. Rolls and A. Treves. The relative advantages of sparse versus distributed encoding for associative neuronal networks in the brain. *Network: Computation in Neural Systems*, 1(4):407–421, 1990.
- [149] C. A. Rothkopf, D. H. Ballard, and M. M. Hayhoe. Task and context determine where you look. *Journal of Vision*, 7(14):16, 2016.
- [150] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 1986.
- [151] M. Sehgal, M. Zhou, A. Lavi, S. Huang, Y. Zhou, and A. Silva. Memory allocation mechanisms underlie memory linking across time. *Neurobiology of Learning and Memory*, 2018.
- [152] A. K. Seth, J. L. McKinstry, G. M. Edelman, and J. L. Krichmar. Texture discrimination by an autonomous mobile brain-based device with whiskers. *EEE International Conference on Robotics and Automation*, (April):4925–4930, 2005.
- [153] H. Z. Shouval. Clusters of interacting receptors can stabilize synaptic efficacies. *Proceedings of the National Academy of Sciences of the United States of America*, 102(40):14440–5, 2005.
- [154] Y. Singer, Y. Teramoto, B. D. B. Willmore, A. J. King, J. W. H. Schnupp, and N. S. Harper. Sensory cortex is optimised for prediction of future input. *bioRxiv*, 2017.
- [155] P. J. Sjöström, A. Rancz, A. Roth, M. Häusser, P. J. Sjöstrom, E. A. Rancz, A. Roth, M. Häusser, S. PJ, R. EA, R. A, H. M, P. J. Sjöstrom, E. A. Rancz,

- A. Roth, and M. Hausser. Dendritic Excitability and Synaptic Plasticity. *Physiological Reviews*, 88(2):769–840, 2008.
- [156] P. J. Sjöström, G. G. Turrigiano, S. B. Nelson, Jesper Sjostrom, G. G. Turrigiano, S. B. Nelson, P. J. Sjöström, G. G. Turrigiano, S. B. Nelson, Jesper Sjostrom, G. G. Turrigiano, and S. B. Nelson. Rate , Timing , and Cooperativity Jointly Determine Cortical Synaptic Plasticity. *Neuron*, 32(6):1149–1164, 2001.
- [157] W. R. Softky and C. Koch. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 13(1):334–50, 1993.
- [158] F. T. Sommer and T. Wennekers. Associative memory in networks of spiking neurons. *Neural networks : the official journal of the International Neural Network Society*, 14(6-7):825–34, 2001.
- [159] I. Song and R. L. Huganir. Regulation of AMPA receptors during synaptic plasticity. *Trends in Neurosciences*, 25(11):578–588, 2002.
- [160] P. O. F. Spie, T. H. E. International, S. For, O. Engineering, C. Koch, and N. Tsuchiya. Relationship between selective visual attention and visual consciousness. *conference paper*, (April), 2016.
- [161] L. Spillmann, B. Dresplangley, and C.-h. Tseng. Beyond the classical receptive field : The effect of contextual stimuli. *Journal of Vision*, 15(9):1–23, 2015.
- [162] G. J. Stuart and N. Spruston. Dendritic integration: 60 years of progress. *Nature Neuroscience*, 18(12):1713–1721, 2015.
- [163] R. Sun and C. L. C. Giles. Sequence learning: from recognition and prediction to sequential decision making. *IEEE Intelligent Systems*, 16(4):67–70, 2001.

- [164] N. Takahashi, K. Kitamura, N. Matsuo, M. Mayford, M. Kano, N. Matsuki, and Y. Ikegaya. Locally synchronized synaptic inputs. *Science*, 2012.
- [165] J. G. Taylor and S. Coombes. Learning Higher Order Correlations. *Neural Networks*, 6:423–428, 1993.
- [166] S. Teki, B. M. Gu, and W. H. W. H. Meck. The Persistence of Memory : How the Brain Encodes Time in Memory. *Current Opinion in Behavioral Sciences*, 17(August):178–185, 2017.
- [167] A. Tran-Van-Minh, R. D. Cazé, T. Abrahamsson, L. Cathala, B. S. Gutkin, and D. A. DiGregorio. Contribution of sublinear and supralinear dendritic integration to neuronal computations. *Frontiers in Cellular Neuroscience*, 9(March):1–15, 2015.
- [168] B. van Bommel and M. Mikhaylova. Talking to the neighbours: The molecular and physiological mechanisms of clustered synaptic plasticity. *Neuroscience and Biobehavioral Reviews*, 71:352–361, 2016.
- [169] D. C. Van Essen. Corticocortical and thalamocortical information flow in the primate visual system. *Progress in Brain Research*, 149(4):173–185, 2005.
- [170] O. V.M. Testing simple neuron models with dendrites for sparse binary image representation. *Искусственный интеллект*, №2, 2017.
- [171] B. Widrow and M. Hoff. Adaptive switching circuits. *IRE WESCON Convention Record*, 1960.
- [172] D. J. Willshaw, O. P. Buneman, and H. C. Longuet-Higgins. Non-holographic associative memory. *Nature*, 222(5197):960–962, 1969.
- [173] F. Wörgötter and B. Porr. Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms. *Neural Computation*, 17(2):245–319, 2005.

- [174] Y. Wu, G. Wayne, A. Graves, and T. Lillicrap. The Kanerva Machine: A Generative Distributed Memory. *arXiv*, pages 1–16, 2018.
- [175] L. C. Yeung, H. Z. Shouval, B. S. Blais, and L. N. Cooper. Synaptic homeostasis and input selectivity follow from a calcium-dependent plasticity model. *Proceedings of the National Academy of Sciences of the United States of America*, 101(41):14943–8, 2004.
- [176] Zachary C. Lipton. A Critical Review of Recurrent Neural Networks for Sequence Learning. *arXiv preprint*, pages 1–35, 2015.
- [177] F. Zenke, E. J. Agnes, and W. Gerstner. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nature communications*, 6:6922, 2015.